

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 19:07:45 ; Search time 6252 Seconds

(without alignments)  
10731.394 Million cell updates/sec

Title: US-10-520-820-28

Perfect score: 1434

Sequence: 1 atgaagaatgaacgtgcacga.....aacagataaanaagctaa 1434

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:\*

1: gb\_esc1.\*

2: gb\_esc2.\*

3: gb\_esc3.\*

4: gb\_esc4.\*

5: gb\_esc5.\*

6: gb\_esc6.\*

7: gb\_esc7.\*

8: gb\_esc8.\*

9: gb\_esc9.\*

10: gb\_esc10.\*

11: gb\_esc11.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	820	57.2	820	10	CL659093 PRI0133A
2	489.8	34.2	895	10	CL655529 PRI0149D
3	419.4	29.2	682	10	CC871613 NDL.2A24.
4	323	22.5	722	8	DN468758 USDA-PP 1
5	252	17.6	879	1	B2569854 msh2_105.
6	237.2	16.5	682	6	CB065364 EST645045
7	234.4	16.3	593	10	CL678544 PRI0123B
8	178	12.4	1135	9	B2564465 pac82-164
9	159.6	11.1	879	10	AG841316 Oryza sat
10	159	11.1	762	10	AG841316 Oryza sat
11	146.8	10.2	1195	9	B2554342 pac81-60.
12	100.8	7.0	505	2	B333786 B533d03.X
13	96	6.7	226	9	B2648805 OGCAM49TM
14	96	6.7	251	9	B2648805 OGCAM49TM
15	90.4	6.3	764	7	CN817146 HRO4528 G
16	90.4	6.3	692	10	CM332900 104.830.1
17	84	5.9	483	9	BH404405 AG-ND-140
18	73.2	5.1	530	9	BH390803 AG-ND-140
19	66.4	4.6	502	10	AG945142 Drosophila
20	63.8	4.4	1474	9	B2552978 pac81-60
21	54.8	3.8	622	9	CC821631 MBSP55 U
22	54.6	3.8	623	7	CO140047 EST834718

23	54.6	3.8	626	7	CO142807	CO142807 EST837478
24	53	3.7	615	7	CO141185	CO141185 EST835856
25	51	3.6	785	6	CF447053	CF447053 EST683398
26	47.4	3.3	959	9	AF075847	AF075847 AF075847
27	46.4	3.2	878	10	CN942436	CN942436 TCB23.4.H
28	45.8	3.2	502	7	CN898041	CN898041 010701AAZ
29	45.2	3.2	584	5	CA030018	CA030018 HX05N16r
30	45	3.1	621	8	DT002298	DT002298 MCRta1016
31	44.8	3.1	715	8	DR479547	DR479547 W602817.C
32	44.8	3.1	715	8	DR485191	DR485191 W602817.C
33	44.8	3.1	715	8	DR486049	DR486049 W602817.C
34	44.6	3.1	723	8	DR480409	DR480409 W602817.C
35	44.6	3.1	414	11	P957R	AL390661 Leishman
36	44.6	3.1	776	8	CX545158	CX545158 UCRPT01.5
37	43.6	3.0	652	5	B0996249	B0996249 OCG12R01
38	43.6	3.0	728	8	DR67881	DR67881 ZM_BBP008
39	43.6	3.0	726	5	BQ850159	BQ850159 QCB1107
40	43.6	3.0	772	10	CZ002297	CZ002297 OA_BBA005
41	43.4	3.0	634	7	CN910981	CN910981 021119ABM
42	43.2	3.0	453	3	BJ935968	BJ935968 BJ935968
43	42.6	3.0	380	7	CO138482	CO138482 EN16913.5
44	42.6	3.0	430	7	CO282447	CO282447 EK200212
45	42.6	3.0	616	3	BI639334	BI639334 SD21893.5

#### ALIGNMENTS

RESULT 1  
LOCUS CL659093 820 bp DNA linear GSS 09-JUL-2004  
DEFINITION PRI0133A\_G12 - PRI0133A\_B21 (820) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

ACCESSION CL659093  
VERSION CL659093.1 GI:50142804  
KEYWORDS GSS.  
SOURCE Pristionchus pacificus  
ORGANISM Pristionchus pacificus  
Bukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.

REFERENCE 1 (bases 1 to 820)  
AUTHORS Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.  
TITLE Apanb: an Acedb database for the nematode satellite organism Pristionchus pacificus  
JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)  
PUBMED 14681447

COMMENT Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@uebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.  
Seq primer: T7  
Class: fosmid ends.  
Location/Qualifiers  
1. 820  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Mixed stage fosmid library of P. pacificus var. California"  
/note="Vector: pBplfos-5 Fosmid vector"

#### FEATURES

##### source

#### ORIGIN

Query Match 57.2%; Score 820; DB 10; Length 820;  
Best local similarity 100.0%; Pred. No. 3.6e-225;  
Matches 820; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 TTGACGGGCAATTGACGATGACGCGCGCTGAGTAATCTCTGCGGACGTCAGTC 273  
 DB 1 TTGACGGGCAATTGACGATGACGCGCGCTGAGTAATCTCTGCGGACGTCAGTC 60  
 QY 274 AAATGCGACTTGGTTCTGTAACGACGATCCGACCTTACCAATTAACGGGTA 333  
 DB 61 AAATGCGACTTGGTTCTGTAACGACGATCCGACCTTACCAATTAACGGGTA 120  
 QY 334 CGGAAACCAACGATCCGCTGCTGTAAGAGGTTTGAAGAGGTTTGAATCCGAG 393  
 DB 121 CGGAAACCAACGATCCGCTGCTGTAAGAGGTTTGAAGAGGTTTGAATCCGAG 180  
 QY 394 CCGCTGACGACGCGGATTAATACGCGCTGAGTTGATGCGCGCTGCTTCTGAC 453  
 DB 181 CCGCTGACGACGCGGATTAATACGCGCTGAGTTGATGCGCGCTGCTTCTGAC 240  
 QY 454 TACGCCAAAGGTGCGCTGCGCAAGGTAAGATGATCCAACTGGCGCTTAAAGCGG 513  
 DB 241 TACGCCAAAGGTGCGCTGCGCAAGGTAAGATGATCCAACTGGCGCTTAAAGCGG 300  
 QY 514 GTTCGGGTGCTGATGATCCAAAGGTAACGATTTTGAAGGCTACGCGCGCTACGCTG 573  
 DB 301 GTTCGGGTGCTGATGATCCAAAGGTAACGATTTTGAAGGCTACGCGCGCTACGCTG 360  
 QY 574 TTAAACGCGAATCTCTCGAATTTGAAGCTGTTGCTGTAATGTAAGACGGAAGAG 633  
 DB 361 TTAAACGCGAATCTCTCGAATTTGAAGCTGTTGCTGTAATGTAAGACGGAAGAG 420  
 QY 634 ATTGTTGAGCGCGGCAATGTAATGCTGATGCTGATGCTGCTGCTGCTGCTGCTG 693  
 DB 421 ATTGTTGAGCGCGGCAATGTAATGCTGATGCTGATGCTGCTGCTGCTGCTGCTG 480  
 QY 694 CGTTCCGAAACAGGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753  
 DB 481 CGTTCCGAAACAGGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 QY 754 CAACGCGAAGAGTGAATGACGTTACCGGTGCGGCGCAACGATGATGCTGCTGCTG 813  
 DB 541 CAACGCGAAGAGTGAATGACGTTACCGGTGCGGCGCAACGATGATGCTGCTGCTG 600  
 QY 814 GCAACGCTGCGAGCGGCTAATTCGCTGGAAGAACCTGCTTCTTGGCAATGCGCGGCT 873  
 DB 601 GCAACGCTGCGAGCGGCTAATTCGCTGGAAGAACCTGCTTCTTGGCAATGCGCGGCT 660  
 QY 874 GGCGTGGTGGTGGGCAACTGGAACCTCAACGCTTTCGCGGATGAGCTGGAATGCT 933  
 DB 661 GGCGTGGTGGTGGGCAACTGGAACCTCAACGCTTTCGCGGATGAGCTGGAATGCT 720  
 QY 934 GTACGTGACGCTGACGATACAGGCTTGGCGTGAATGACCGAAGAGAACTGAAGCTG 993  
 DB 721 GTACGTGACGCTGACGATACAGGCTTGGCGTGAATGACCGAAGAGAACTGAAGCTG 780  
 QY 994 GTACGCGACGCGCTTAAACCTGCTGGAAGAAAGTGGTGAATGA 1033  
 DB 781 GTACGCGACGCGCTTAAACCTGCTGGAAGAAAGTGGTGAATGA 820

RESULT 2  
 CL665529/c 895 bp DNA linear GSS 09-JUL-2004  
 LOCUS PR01019d.H06 - PR01019d.B21 (895) Mixed stage fosmid library of P.  
 pacificus var. California Pristionchus pacificus genomic, genomic  
 survey sequence.  
 ACCESSION CL665529  
 VERSION CL665529.1 GI:50157226  
 KEYWORDS GSS.  
 SOURCE Pristionchus pacificus  
 ORGANISM Pristionchus pacificus  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
 Neodiplogasteridae; Pristionchus.  
 REFERENCE 1 (bases 1 to 895)  
 AUTHORS Strihsaen,U., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.U.  
 TITLE Appadbi: an Acedb database for the nematode satellite organism

JOURNAL Pristionchus pacificus  
 PUBMED Nucleic Acids Res. 32 (1), D421-D422 (2004)  
 14681447  
 COMMENT Contact: Sommer RJ  
 Evolutionary Biology  
 Max-Planck-Institute for Developmental Biology  
 Spemannstr. 37-39, Tuebingen D-72076, Germany  
 Tel: 00497071601371  
 Fax: 00497071601498  
 Email: ralf.sommer@uebingen.mpg.de  
 This library was generated at Caltech, Pasadena, USA and end  
 sequenced at Vancouver, Canada.  
 Seq primer: T7  
 Class: fosmid ends.  
 Location/Qualifiers  
 1..895  
 /organism="Pristionchus pacificus"  
 /mol\_type="genomic DNA"  
 /strain="California"  
 /db\_xref="taxon:54126"  
 /clone\_lib="Mixed stage fosmid library of P. pacificus  
 var. California"  
 /note="Vector: pBfi08-5 Fosmid vector"

ORIGIN  
 Query Match 34.2%; Score 489.8; DB 10; Length 895;  
 Best Local Similarity 97.5%; Pred. No. 8,7e-130;  
 Matches 508; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 915 GATGAGCTGGAATAATGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 974  
 DB 895 GATGAGCTGGAATAATGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 836  
 QY 975 AGAGAACTGAAGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCT 1033  
 DB 835 AGAGAACTGAAGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCT 776  
 QY 1034 CCAACGCTGCTTGAATCTGCACTGCAACGCGGCAAGCTCTTATCTGCAATGCGGCA 1093  
 DB 775 CCAACGCTGCTTGAATCTGCACTGCAACGCGGCAAGCTCTTATCTGCAATGCGGCA 716  
 QY 1094 AGCTGGTGAACGCTGATGTTGCTGCTGCAACGCAATGCTCTTCAACCAACGCTGAA 1153  
 DB 715 AGCTGGTGAACGCTGATGTTGCTGCTGCAACGCAATGCTCTTCAACCAACGCTGAA 656  
 QY 1154 GGGATCCCGCCCGGTAACCCCACTGGAACAGCGTATGATGCTGGGCGCACTGGAAG 1213  
 DB 655 GGGATCCCGCCCGGTAACCCCACTGGAACAGCGTATGATGCTGGGCGCACTGGAAG 596  
 QY 1214 CGGTGCACTGGTGAATGCTGTTGAAGAGGACACGCGGCAAGCTTGAATGCGCGGATCT 1273  
 DB 595 CGGTGCACTGGTGAATGCTGTTGAAGAGGACACGCGGCAAGCTTGAATGCGCGGATCT 536  
 QY 1274 TGGCAATGCTGCTGTAAGAGGCGGCTATTAACCAAGAGATGCTGGGAGTAAG 1333  
 DB 535 TGGCAATGCTGCTGTAAGAGGCGGCTATTAACCAAGAGATGCTGGGAGTAAG 476  
 QY 1334 AAGTGGGCGCAACGCTGCGGAAGGTTGCTGCTCAACTTGAAGACGCTGCTGACGA 1393  
 DB 475 AAGTGGGCGCAACGCTGCGGAAGGTTGCTGCTCAACTTGAAGACGCTGCTGACGA 416  
 QY 1394 CCAACATCATCAAGAGATCCCAACAGATTAAGAGCTTAA 1434  
 DB 415 CCAACATCATCAAGAGATCCCAACAGATTAAGAGCTTAA 375

RESULT 3  
 CC871613 682 bp DNA linear GSS 24-JUL-2003  
 LOCUS ND1.2A24.spe Notre Dame Liverpool Aedes aegypti genomic clone  
 DEFINITION NotreDame Liverpool-2A24, genomic survey sequence.  
 ACCESSION CC871613  
 VERSION CC871613.1 GI:33231623

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

G99.  
Aedes aegypti (yellow fever mosquito)  
Aedes aegypti  
Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
Culicidae; Culicini; Aedes; Stegomyia.  
1 (bases 1 to 682)  
Loftus, B., Shetty, J., Knudson, D. and Severson, D.  
BAC end sequencing of Aedes aegypti  
Unpublished (2003)  
Other GS88: NDJ.2A24.T7  
Contact: Brendan Loftus  
Department of Bacteriologic Genomics  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: entoe@tigr.org  
Library was provided by David Severson  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
Location/Qualifiers  
1..682  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="Liverpool"  
/db\_xref="taxon:7159"  
/clone="Notredame Liverpool-2A24"  
/clone\_lib="Notre Dame Liverpool"  
/note="Vector: pECBAC1, Site 1: Hind III; The library was  
prepared from whole body tissue of newly hatched L1 larvae  
by David Severson at the University of Notre Dame and  
Hongbin Zhang"

ORIGIN  
Query Match 29.2%; Score 419.4; DB 9; Length 682;  
Best Local Similarity 76.2%; Pred. No. 1.8e-109;  
Matches 516; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 286 GTTTCGTGACGAGCATCCGACATTCACCAATTCAGGGTACTTTCGCCGACCAACAG 345  
Db 1 GTCTCGGTCGCGACCCATCCGACATCCACCAAGCTGCGCGTCTGCGCGACACAG 60  
QY 346 CTGATCCGTCTGGAATTTGAAGAGCTTTCGAAGTGTATTCGCGCGCGCTGACAG 405  
Db 61 CTGATCCGTCTGGAATTTGAAGAGCTTTCGAAGTGTATTCGCGCGCGCTGACAG 120  
QY 406 CGGATTCAGCGCGCTGAGTTGAGGCGCGCTGCTTTCGACTACGCCAAGT 465  
Db 121 CGGATTCAGCGCGCTGAGTTGAGGCGCGCTGCTTTCGACTACGCCAAGT 180  
QY 466 GCGCTGCGACGCGTACGACAGATGATCCACTGCGCGGTAAAGCGGTGTTCCGCTG 525  
Db 181 GCGCTGCGACGCGTACGACAGATGATCCACTGCGCGGTAAAGCGGTGTTCCGCTG 240  
QY 526 ATTATTCACCAAGTACCGATTTGAGGCGCTACCGCGCGCTACGCTGTAAACCCGAT 585  
Db 241 ATCATCTTAAAGCTCTGACTTGAGCGCTATGCGCGCGAGCGCTGACGCCCAAC 300  
QY 586 CTCTCGGAATTTGAAGCTGTGTGCGTAAATGTAAAGACCGAAGAAAGATGTTGAGCGC 645  
Db 301 CTCTCGGAATTTGAAGCTGTGTGCGTAAATGTAAAGACCGAAGAAAGATGTTGAGCGC 360  
QY 646 GCGATGAACTGATTCGATTCGAACTCTCGGCTCTGTTAGTACCCGTTCCGAAACG 705  
Db 361 GCGATGAACTGATTCGATTCGAACTCTCGGCTCTGTTAGTACCCGTTCCGAAACG 420  
QY 706 GGTATTCGCTGCTGCAACCGGCTAAAGCGCGCTGCAATAGCCAAACGACGACAGAA 765  
Db 421 GGTATTCGCTGCTGCAACCGGCTAAAGCGCGCTGCAATAGCCAAACGACGACAGAA 480  
QY 766 GTGATGACGTTACCGGTCGCGGCGACACGCTGATTCGCTGCGCGCAACGCTGCA 825  
Db 825 GTGATGACGTTACCGGTCGCGGCGACACGCTGATTCGCTGCGCGCAACGCTGCA

Db 481 GTGTTGATGATGACCGGCGCGCGACACGCTGATGCGCGTCTGCGCGCTGCGCC 540  
QY 826 GCGGTAATTCGCTGGAAGAGCTCTTCTTTCGCAATGCGCGCGCTGCGCTGCTG 885  
Db 541 GCGGTAATTCGCTGGAAGAGCTCTTCTTTCGCAATGCGCGCGCTGCGCTGCTG 600  
QY 886 GCGAATTCGGAAGCTCTGACGCTTTCGCGCGCTGCAATGCGCGCGCTGCAATG 945  
Db 601 GCGAATTCGGAAGCTCTGACGCTTTCGCGCGCTGCAATGCGCGCGCTGCAATG 660  
QY 946 GCAGATACAGGCTTTG 962  
Db 661 GCCGAAACCGGCTTTG 677

RESULT 4  
DN468758/c 722 bp mRNA linear EST 01-AUG-2005  
LOCUS  
DEFINITION  
Diaphorina citri CDNA clone WHDc040\_E04 5', mRNA sequence.  
ACCESSION  
DN468758.1 GI:71538215  
VERSION  
DN468758.1  
KEYWORDS  
EST.  
Diaphorina citri  
Diaphorina citri  
ORGANISM  
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Psylliformes;  
Psyllidae; Psyllidae; Diaphorina.  
REFERENCE  
1 (bases 1 to 722)  
Hunter, W.B., Dang, P.M. and McKenzie, C.L.  
Gene expression in adult Asian citrus psylla, Diaphorina citri  
Kuwayama (Hemiptera: Psyllidae)  
Unpublished (2006)  
JOURNAL  
COMMENT  
U.S. Horticultural Research Phat Dang, USDA, ARS  
U.S. Horticultural Research Lab, Subtropical Insect Research Unit  
USDA-ARS  
2001 South Rock Road, Ft. Pierce, FL 34945, USA  
Tel: (772) 462-5988, (772) 462-5940  
Fax: (772) 462-5986  
Email: Whunter@usnrl.ars.usda.gov  
Seq primer: T3 primer.  
Location/Qualifiers  
1..722  
/organism="Diaphorina citri"  
/mol\_type="mRNA"  
/db\_xref="taxon:121845"  
/clone="WHDc040\_E04"  
/sex="Mixed genders"  
/tissue\_type="whole body"  
/dev\_stage="adults"  
/lab\_host="X11-Blue"  
/clone\_lib="Diaphorina citri Kuwayama (Hemiptera:  
Psyllidae)"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; Source: Colonies maintained by Dr. McKenzie, USNRL,  
Fort Pierce, FL. A high quality EST with at least 200  
contiguous bases at Trace Turner score of 20 or better.  
Library construction by Laura Hunsicutt, USDA, ARS, U.S.  
Horticultural Res. Lab, Ft. Pierce, FL, USA."

ORIGIN  
Query Match 22.5%; Score 323; DB 8; Length 722;  
Best Local Similarity 100.0%; Pred. No. 1.2e-81;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 CCGACATTCACCAATTCAGGTAATTCGCGACCAACAGCTGATTCGCTGATTTT 363  
Db 722 CCGACATTCACCAATTCAGGTAATTCGCGACCAACAGCTGATTCGCTGATTTT 663  
QY 364 GAAGAAAGTTTGAAGGTTGATTCGCGACGCGCTGCAAGCGGATTAATCAGGCTG 423  
Db 662 GAAGAAAGTTTGAAGGTTGATTCGCGACGCGCTGCAAGCGGATTAATCAGGCTG 603

```

Qy 424 AGTTCGATGGCGCGCTGCTGCTTCTGATCACTACCCAAAGGTGCGCTGCGAAGCGTACAG 483
Db 602 AGTTGATGGCGCGCTGCTGCTTCTGATCACTACCCAAAGGTGCGCTGCGAAGGTACAG 543
Qy 484 CAGATGATCCAACTGCGCGCTGTAAGCGGGTGTTCGGTGTCTGATGATTCAGAAAGGTACC 543
Db 542 CAGATGATCCAACTGCGCGCTGTAAGCGGGTGTTCGGTGTCTGATGATTCAGAAAGGTACC 483
Qy 544 GATTTAGAGCGCTACCGCGCGCTACGCTGTTAAAGCGCAATCTCTGGAAATTTGAAGCT 603
Db 482 GATTTAGAGCGCTACCGCGCGCTACGCTGTTAAAGCGCAATCTCTGGAAATTTGAAGCT 423
Qy 604 GTTGTCCGTAATATGTAAGCCGA 626
Db 422 GTTGTCCGTAATATGTAAGCCGA 400

RESULT 5
BZ569854 879 bp DNA linear GSS 17-DEC-2002
LOCUS meh2.105.x2.meh Pseudomonas aeruginosa genomic clone meh2.105,
DEFINITION genomic survey sequence.
ACCESSION BZ569854
VERSION BZ569854.1 GI:27204915
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 879)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Class: shotgun.
FEATURES
source Location/Qualifiers
1..879
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/db_xref="taxon:287"
/clone="meh2.105"
/clone_1kb="meh"
/notice="Environmental isolate. Whole genomic shotgun
library."
ORIGIN
Query Match 17.6%; Score 252; DB 9; Length 879;
Best Local Similarity 59.5%; Pred. No. 4,2e-61;
Matches 454; Conservative 0; Mismatches 307; Indels 2; Gaps 2;
Qy 542 CCGATTGTTAGCGCTACCGCGCGCTACGCTGTTAAGCCGAATCTCTCGGAATTTGAAG 601
Db 91 CCGACTTCGCATCTATCGCGCGCGCGCTGATCACCCGAACTGTCCGAATTCAGA 150
Qy 602 CTGTGTCGCTAATATGTAAGCCGAAGAGATTTGAGCGGGCATGAATGATG 661
Db 151 CCATGCTGCGCGCTGCGCGCGCGCAAGGCGAATGCTGCGCAAGGCGCGCTGATGA 210
Qy 662 CCGATTACGAATCTCGGCTCTGTTAGTGAACCGCTTCGAAACAGGTAATGCTGCTGC 721
Db 211 GCGAACTGACCTGCTGCTGCTGCTGATGACCGCGCGGAGCATGACATGACCTGCTCC 270
Qy 722 AACCGGTAAGCGCGCTGCTGATGACCAACCGCAAGCGCAAGAGTATGACCTTACCG 781

```

```

Db 271 GCCATGACCAAGCGCGCTGACCTTCGCGCGCGCGCGGAGATGTTGACGTACCG 330
Qy 782 GTCGCGGCGCAACAGGATTTGCGCTCTGCGCGCAACGCTGACGGCGGATATTCGCTGG 841
Db 331 GTTCGCGCGATACAGGATCTTCACTTCCGCGCGCGCTTCCGCGCGGAGAGCTTC 330
Qy 842 AAGAACCTGCTCTTCTTGCATATGCGCGCGCTGCGCTGCTGCGCAACTGGAACCT 901
Db 391 CTTCGCGCGGATGCTGCGCAACCTGCGCGCGCGGATGCTGCGCAAGCTGGATACCG 450
Qy 902 CACGCGTTTCGCGCGATGACGTGGAATAATGCTGTACTGTGACGTGACATGAGCTTGG 951
Db 451 CGCGATTCAGCGCGCGCGCACTGCGCGCGGTGAGCGCGAGCAAGGTTCCGACGCTG 510
Qy 962 GCGTATGACCGAAGAGAACTGAAGCTGCGCGCTAGCGGACGCGCTAAACGTGTGAAA 1021
Db 511 GCGTCTGCGCTGAGAGCAATTTCTCTGCAATCAAGACGCGCGCGCGCGAGCA 570
Qy 1022 AAGTGTGATGACCAACGCTGCTTGTGACATCTCTGACGCGCGGCAAGTCTTATCTGG 1081
Db 571 AGATGCTTCAACCAATGAGGTGCTTGAATCTTCAACGCGCGCGCAAGTACTTACCTCG 630
Qy 1082 CAATGCGCGCAAGCTGCGTGAACGCTTGAATTTGCGCTGCAACAGCATGCTTCAACCA 1141
Db 631 AACAGCGCGCGCGCGCGCGAGCGCGCTGATGCGAGGTCAGCAACGAGCGCTTGGTCA 690
Qy 1142 AACGCGTGAAGGAGGATTCGCGCGCGCGGTAAACCACTGCAACGCGTATGATGTGCTGG 1201
Db 691 CTGCGCTGAAGGCGCTTGGCGCGCGCAT-AATCTGCGTGAACGCGCGCGCTGCGTACTCG 749
Qy 1202 GCGCACTGGAAGCGGTGACACTGCGTGAATGCTGTTGAAGAGACAGCGCGCGCTTGA 1261
Db 750 CCGGCGCT-GNNCGGTGAGACTGAGTGTGAGGCTTCCGCCAAACACTCTGAAAGCGCTTG 808
Qy 1262 TCGCCGGGATCTTCCAGATCTGCTGTGTAAGAGCGCGCACTA 1304
Db 809 TTGAGCAGGTGCGCTCGAGCGTGTGACAGGCGCGCGAATA 851

RESULT 6
CB065364 682 bp mRNA linear EST 21-JAN-2003
LOCUS EST645045 HOGA Medicago truncatula cDNA clone HOGA-211, mRNA
DEFINITION sequence.
ACCESSION CB065364
VERSION CB065364.1 GI:27810942
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
REFERENCE 1 (bases 1 to 682)
AUTHORS Hahn,M.G., Ojane-Reuhs,T., Samad,D., Town,C.D., Van Aken,S.,
Utecherback,T., Cho,J. and Fraser,C.M.
TITLE ESTs from roots of Medicago truncatula treated with
oligogalacturonides of DP 6-20
JOURNAL Unpublished (2001)
COMMENT Contact: Michael G. Hahn
Complex Carbohydrate Research Center
University of Georgia
220 Riverbend Road, Athens, GA 30602-4712, USA
Tel: 706-542-4457
Fax: 706-542-4412
Email: hahn@ccrc.uga.edu
TIgR sequence name: MTMAE49TK
More information is available at: www.medicago.org
Seq primer: Skmod (CTA gaa cta gtc gat cc).
FEATURES
source Location/Qualifiers
1..682
/organism="Medicago truncatula"
/mol_type="mRNA"

```



```

/cultivar="A17"
/db.xref="taxon:3860"
/clone="HOGA-211"
/tissue_type="3 day old seedling roots"
/dev_stage="24 hours after treatment in the dark at 26 C
with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
presence of 100 ug/ml Gentamicin"
/lab_host="XL0LR"
/clone_idb="HOGA"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the UniZap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were packaged
from the recombinant lambda-Zap phage using Ex-asebist
helper phage and propagated in SOLR cells."

```

## ORIGIN

Query Match	16.5%	Score 237.2;	DB 6;	Length 682;
Best Local Similarity	59.5%;	Pred. No. 7.4e-57;		
Matches 401;	Conservative 0;	Mismatches 273;	Indels 0;	Gaps 0;

Qy	577	ACGCCAATCTCTCGGAATTGGAACCTGTGTGCGTAAATGTAACCGAAGAAAGATT	636
Db	2	ACTCCGAACCTCAGGGAATTGMAAACATCGTCCGTGCTGGCGCTGACGAGAACGATTTG	61
Qy	637	GTTGAGCGCGGCATGAACTGATATGCCAATTACGAATCTTCGGCTCTGTATGTAACCGT	696
Db	62	GTCCCAAGGGCCCTTCAACTGCTGCGAGATCTCGACTCGGGCGCACTCGGTGACCCGC	121
Qy	697	TCCGAACAGGGTATGTGCTGCTGCAACCGGGTAAAGGCGCGCTGCATATGCAACCCCA	756
Db	122	GCGCAATGAGCATGACCTCTGTGTGACCGGTCAACCGCAATGCACTTGCCGCGCGG	181
Qy	757	GCGCAGAAAGTATGATGACTTACCGGTGCGGGCGACACAGTGAATTTGGCGTCTTGCGGCA	816
Db	182	GCCGTGAAAGTTCATGATGTCACCGGTGCGCGGCATACCTGATCTCAACCTCGCCGG	241
Qy	817	ACGCTGGCAGCGGGGTAATTCCGTGGAAGAAAGCTGCTCTTTTGCMAATGCGCGCGCTGC	876
Db	242	GCCATTCGGGCTGGGAGAACTTCGCCCATGCTCGTGGCTTGCGCCAACCTGGCTGCGGA	301
Qy	877	GTGGTGTGCGGCMAACTGGGAACCTCCACGGTTTGCAGATGACCTGAAATACCTGTA	936
Db	302	ATCGTGTGCGCAAACTGGATACGGCGGCATCAAGCGCCGGAAATTGGCCCGCGATC	361
Qy	937	CTGTGACGTGCAGATACAGGCTTTTGGCGTATGACCGAAGAGAACTGAAGCTGGCCGTA	996
Db	362	CAGCGCGAAGAGGGGCTTCGAGCGCGGGGTGCTGGAGCACTCTGCTGCGCATC	421
Qy	997	GCGGCAGCGCTTAAACGTGTGTAAGAAAGTGTGATGACCAACCGGTCTTTGACATCCG	1056
Db	422	GACGATGCGCGCGGCACACGAGAAAGATCGTGTCAACATGTTGCTTGACATCTTC	481
Qy	1057	CAGCGCGGGCAGCTCTTATCTTGCAAAATGCCCGCAAGCTGGGTGAACCGCTTGATTTG	1116
Db	482	CAGCCCGGGCAGTAACTTACTTGAGCAAGCTCTGTGCCAAGGGAGATCGCTGATGTC	541
Qy	1117	GCCGTCAACACGATGCTTCCACCAAAAGGCTGAAAGGGAGATCCCGCCCGTAAACCA	1176
Db	542	GCGGTCAACGACGAGCTTCCGTCAAGCGGCTCAAGAGGACCAAGTGTGCCATCAACAGT	601
Qy	1177	CTCGAACAGCTATGATTTGTGCTGGGCGCACTGAAAGCGTGCATGTGGTATGATGCTTT	1233
Db	602	GTGATGCGGCATATGCGCTGTGCTGTCTGGAGAGCGGTGGACTGGGTATCAACTTC	661
Qy	1237	GAAGAGACAGCC 1250	
Db	662	CCGAGACACCCC 675	

RESULT 7  
CL678544/C

LOCUS	Ct678544	593 bp	DNA	linear	GSS 09-JUL-2002
DEFINITION	PRI0123b_A02_2 - PRI0123b_BR (593) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.				
ACCESSION	Ct678544				
VERSION	Ct678544.1	GI:50184914			
KEYWORDS	GSS.				
SOURCE	Pristionchus pacificus				
ORGANISM	Pristionchus pacificus Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus. 1 (bases 1 to 593) Srinivasan,J., Otto,G.W., Kahlow,U., Gislser,R. and Sommer,R.J. Appends: an Acedb database for the nematode satellite organism Pristionchus pacificus Nucleic Acids Res. 32 (1), D421-D422 (2004)				
JOURNAL	14681447				
COMMENT	Contact: Sommer RJ Evolutionary Biology Max-Planck-Institute for Developmental Biology Spemannstr. 37-39, Tuebingen D-72076, Germany Tel.: 00497071601371 Fax: 00497071601498 Email: ralf.sommer@uebingen.mpg.de This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada. Seq primer: T7 Class: fosmid ends. Location/Qualifiers 1..593 /organism="Pristionchus pacificus" /mol_type="genomic DNA" /strain="California" /db_xref="taxon:54126" /clone_lib="Mixed stage fosmid library of P. pacificus var. California" /note="Vector: pspifos-5 Fosmid vector"				
ORIGIN					
Query Match	16.3%; Score 234.4; DB 10; Length 593;				
Best Local Similarity	99.2%; Pred. No. 4.6e-56;				
Matches 246; Conservative 0; Mismatches 1; Indels 1; Gaps 1.					
OY	1188 TATGATTGTGCT-GGGCGCACTGGAAGCGTGCAGCTGGTGATGTTCGTTGAAGAAGCA	1246			
DB	593 TATGATTGTGCTGGGGCGCACA CTGGAACGGTGCAGCTGGTGATGTTCGTTGAAGAAGCA	534			
OY	1247 CGCCGACCGCTTATCGCCGGGATCTTGCAGATCTGCTGTAAAGGCGCGCATATA	1306			
DB	533 CGCCGACCGCTTATCGCCGGGATCTTGCAGATCTGCTGTAAAGGCGCGCATATA	474			
OY	1307 AACAGAGAAGATTCCGGGAGTAAGAAAGTCTGGGCAACGTCGCGAATGTTCGGTGC	1366			
DB	473 AACAGAGAAGATTCCGGGAGTAAGAAAGTCTGGGCAACGTCGCGAATGTTCGGTGC	414			
OY	1367 TCAACTTTGAAGCGTTGCTCGACGACCAATCATCAAGAAATCCAACGATATAA	1426			
DB	413 TCAACTTTGAAGCGTTGCTCGACGACCAATCATCAAGAAATCCAACGATATAA	354			
OY	1427 AAGGCTAA 1434				
DB	353 AAGGCTAA 346				
RESULT 8					
BZ564465/c	1135 bp	DNA	linear	GSS 17-DEC-2002	
LOCUS	BZ564465				
DEFINITION	pacer2-164_4762.v2 pacer2-164 Pseudomonas aeruginosa genomic clone				
ACCESSION	BZ564465				
VERSION	BZ564465.1	GI:27190368			
KEYWORDS	GSS.				
SOURCE	Pseudomonas aeruginosa				

ORGANISM *Pseudomonas aeruginosa*  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 1135)  
 AUTHORS Spencer,D.H., Raymond,C.K., Smith,B.E., Sims,B.E., Hastings,M.,  
 Burns,J.L., Kaul,R. and Olsen,M.V.  
 TITLE Whole-Genome-Sequence variation among multiple isolates of  
*Pseudomonas aeruginosa* library  
 JOURNAL J. Bacteriol. (2002) In press  
 COMMENT Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: ckraymond@u.washington.edu  
 Class: Shotgun.  
 Location/Qualifiers  
 source 1..1135  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="2-164"  
 /db\_xref="taxon:287"  
 /clone="pac82-164\_4762"  
 /clone\_id="pac82-164"  
 /note="clinical isolate 2-164 Whole genomic shotgun  
 library."

FEATURES  
 source 1..1135  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="2-164"  
 /db\_xref="taxon:287"  
 /clone="pac82-164\_4762"  
 /clone\_id="pac82-164"  
 /note="clinical isolate 2-164 Whole genomic shotgun  
 library."

ORIGIN  
 Query Match 12.4%; Score 178; DB 9; Length 1135;  
 Best Local Similarity 57.7%; Pred. No. 1.1e-39;  
 Matches 336; Conservative 0; Mismatches 245; Indels 1; Gaps 1;

438 GCTGCTGCTTCTTCTGATACGCGCAAGGCGCTGGCA-AGCGTACGACAGATGATCCAC 496  
 DB GGTGCTGCTTCTGATACGCGCAAGGCGCGCGCAAGACCTCCGCTGATCCACG 610  
 497 TGGCGCGTAAAGCGGCTGTCGCGTGTGATTCGCAAGTACGATTTGAGCGCT 556  
 DB CGGTGCGGCGCGCGCAATTCGCGTATCTGCGCGCATCCAGGCGCAAGTCTCGCATCT 550  
 557 ACCGCGCGCTACGCTGTTAAGCGCAATCTCTCGCAATTTGAAGCTGTGTCTGTAAT 616  
 DB ATCCGCGCGCGCGCGCTGATCAACCGCAACCTGCGCAATTCGCAATCTCTGCGCTT 490  
 549 ATCCGCGCGCGCGCGCTGATCAACCGCAACCTGCGCAATTCGCAATCTCTGCGCTT 490  
 617 GTAAGACCGAAGAGATTTGTAAGCGCGCGCATGAAATGATTCGCGATTAAGAACTCT 676  
 DB GCGCGCAAGAGCGCAATGATTCGCGCAAGGCGCGCGCATGATGAGCAACTGCACTCG 430  
 489 GCGCGCAAGAGCGCAATGATTCGCGCAAGGCGCGCGCATGATGAGCAACTGCACTCG 430  
 677 CGGCTCTGTATGATACCGGTTCCGAACAGGGTATGCTGCTGCAACCGGGTTAAAGCGC 736  
 DB GTGCTCTGTGTATGATACCGGTTCCGAACAGGGTATGCTGCTGCAACCGGGTTAAAGCGC 370  
 429 GTGCTCTGTGTATGATACCGGTTCCGAACAGGGTATGCTGCTGCAACCGGGTTAAAGCGC 370  
 737 CGCTGATATATCCAAACCGAAGCGGAGATGATGATGATGATGATGATGATGATGATGAT 796  
 DB CGCTGATATATCCAAACCGAAGCGGAGATGATGATGATGATGATGATGATGATGATGAT 796  
 369 CCCTGCAACTCCGCGCGCGCGCGCGGAGATGATGATGATGATGATGATGATGATGATGAT 310  
 DB CCCTGCAACTCCGCGCGCGCGCGCGGAGATGATGATGATGATGATGATGATGATGATGAT 310  
 797 TGATTTGGGCTCTGCGCGCGCAACGCTGGAAGGGGTAATTCGCTGGAAGGAAGCGCTTCT 856  
 DB TGATTTGGGCTCTGCGCGCGCGCAACGCTGGAAGGGGTAATTCGCTGGAAGGAAGCGCTTCT 856  
 309 TCATCTCCACCTGCG 250  
 DB TCATCTCCACCTGCG 250  
 857 TTGCGCAATGCG 916  
 DB TTGCGCAATGCG 916  
 249 TGGCGCAACTGCG 190  
 DB TGGCGCAACTGCG 190  
 917 TCGAGCTGGAAGATGCTGTACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGAC 976  
 DB TCGAGCTGGAAGATGCTGTACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGAC 976  
 189 CCGAAGCTGCTGCG 130  
 DB CCGAAGCTGCTGCG 130  
 977 AGGAATGGAAGCTGCG 1018  
 DB AGGAATGGAAGCTGCG 1018  
 129 ACCAATGCTGTGCGCAATCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 88  
 DB ACCAATGCTGTGCGCAATCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 88

RESULT 9  
 AG841316 879 bp DNA 1linear GSS 03-NOV-2004  
 LOCUS AG841316  
 DEFINITION Oryza sativa (indica cultivar-group) genomic DNA, BAC end sequence,  
 BAC clone:K0061H04\_F, genomic survey sequence.  
 ACCESSION AG841316  
 VERSION AG841316.1 GI:55307551  
 KEYWORDS GSS.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)  
 BUKARYOTA; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1  
 AUTHORS Katagiri,S., Wu,J., Ito,Y., Karasawa,W., Shibata,M., Kanamori,H.,  
 Katayose,Y., Nemiki,N., Matsumoto,T. and Sasaki,T.  
 TITLE End Sequencing and Chromosomal In situ Mapping of BAC Clones  
 Derived from an Indica Rice Cultivar, Kasalath  
 Breeding Science 54, 273-279 (2004)  
 2 (bases 1 to 879)  
 Sasaki,T., Matsumoto,T. and Wu,J.  
 Direct Submission  
 Submitted (29-Oct-2004) Takuji Sasaki, National Institute of  
 Agricultural Sciences, Rice Genome Research Program, Kannondai  
 2-1-2, Tsukuba, Ibaraki, 305-8602, Japan  
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 The orientation of the sequence is from T7 side of the BAC clone.  
 Location/Qualifiers  
 source 1..879  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="Kasalath"  
 /db\_xref="taxon:39946"  
 /clone="K0061H04\_F"

ORIGIN  
 Query Match 11.1%; Score 159.6; DB 10; Length 879;  
 Best Local Similarity 56.6%; Pred. No. 2e-34;  
 Matches 316; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

356 TGGATTTTGAAGAGGTTTCGAAGGTGTGATCCGACCGCGCTGCAAGCGGATTAATC 415  
 DB TCGATTTTCGAGAGAGCTCCACCGATACCGTGTGCGCGCAAGGCTCAGCAATTCATTA 61  
 416 AGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 475  
 DB CCCTGTGCGCGCGCACTGACAGCTGATCTCTTCTCGCACTGACCGCAAGGCGCGCTGCA 121  
 62 CCCTGTGCGCGCGCACTGACAGCTGATCTCTTCTCGCACTGACCGCAAGGCGCGCTGCA 121  
 476 GCGTACAGCAGATGATCCAACTGCGCGCGTAAAGCGGCTGTCGCGTGTATGATCCAA 535  
 DB ATGTGCGCGAGATGATGCGCACCGCGCGCGAGCTCGGCAAGCGCATCTGTGATCCCA 181  
 122 ATGTGCGCGAGATGATGCGCACCGCGCGCGAGCTCGGCAAGCGCATCTGTGATCCCA 181  
 536 AAGGTACGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAGCGCAATCTCTCGAAT 595  
 DB AAGGTACGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAGCGCAATCTCTCGAAT 595  
 182 AAGGCGATGATTTTCAAGCGCTATGTGGCGCATGTTGCTGACGCCCAAGATCCGAGC 241  
 DB AAGGCGATGATTTTCAAGCGCTATGTGGCGCATGTTGCTGACGCCCAAGATCCGAGC 241  
 596 TTGAAGCTGTGTCGCTGTAATGTAAGACGGAAGAAAGATGTTGAAGCGCGGATGAAC 655  
 DB TTGAAGCTGTGTCGCTGTAATGTAAGACGGAAGAAAGATGTTGAAGCGCGGATGAAC 655  
 242 TGGTGGTATGTCGCGCGCGAGTGAAGAACGAGACATTCACCGCGCGCGCGCGCGCGCG 301  
 DB TGGTGGTATGTCGCGCGCGAGTGAAGAACGAGACATTCACCGCGCGCGCGCGCGCGCGCG 301  
 656 TGAATGCGGATTTGCAACTCTCGGCTCTGTTAGTGAACCGGTTCCGAACAGGGTATGTCG 715  
 DB TGAATGCGGATTTGCAACTCTCGGCTCTGTTAGTGAACCGGTTCCGAACAGGGTATGTCG 715  
 302 TGGCGCGCGAGCTGCACTGCAAGCGCTGTGCTGACCGGCTGGAAGAGGAGATGAGCC 361  
 DB TGGCGCGCGAGCTGCACTGCAAGCGCTGTGCTGACCGGCTGGAAGAGGAGATGAGCC 361  
 716 TGTGCAACCGGGTAAAGCGCGCTGCATATGCAACCGAAGCGCAAGAAATGTATGACG 775  
 DB TGTGCAACCGGGTAAAGCGCGCTGCATATGCAACCGAAGCGCAAGAAATGTATGACG 775  
 362 TGTACGCGATGCGAGGTTAC---GCACTTCCCGACATAGCGCGCGCGAGGTGTCGATG 418  
 DB TGTACGCGATGCGAGGTTAC---GCACTTCCCGACATAGCGCGCGCGAGGTGTCGATG 418  
 776 TTACCGGTGCGCGCGCACCGGTATGTCGCTCTGCGCGCAACGCTGCGCGCGGTAAT 835  
 DB TTACCGGTGCGCGCGCACCGGTATGTCGCTCTGCGCGCAACGCTGCGCGCGGTAAT 835  
 419 TCTCGGTGCGCGCGATACCGTATCGGACGCTGCGCGCGGATGTCGCGCGCGCGCGCGCAT 478  
 DB TCTCGGTGCGCGCGATACCGTATCGGACGCTGCGCGCGGATGTCGCGCGCGCGCGCGCAT 478

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BZ554342	BZ554342	GI:27161567		
ORGANISM	Pseudomonas aeruginosa			
SOURCE	Pseudomonas aeruginosa			
KEYWORDS	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.			
AUTHORS	Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.			
TITLE	Whole-Genome-Sequence Variation among multiple isolates of Pseudomonas aeruginosa library			
JOURNAL	J. Bacteriol. (2002) In press			
COMMENT	Contact: Chris K. Raymond			
	Genome Center			
	University of Washington			
	Box 352145, Seattle, WA 98105-2145, USA			
	Tel: 2062216954			
	Fax: 206657244			
	Email: ckraymond@u.washington.edu			
FEATURES	Class: Shocgun.			
source	Location/Qualifiers			
	1..1195			
	/organism="Pseudomonas aeruginosa"			
	/mol_type="genomic DNA"			
	/strain="1-60"			
	/db_xref="taxon:287"			
	/clone="pac61-60_464"			
	/clone_11b="pac61-60"			
	/notes="Clinical isolate 1-60 whole genomic shotgun library."			
ORIGIN				
Query Match	10.2%; Score 146.8; DB 9; Length 1195;			
Best Local Similarity	60.4%; Pred. No. 1.1e-30;			
Matches 241; Conservative	0; Mismatches 158; Indels 0; Gaps 0;			
Db	539 GTACCAATTTGAGCGCTACCGCGCGCTACGCTGTTAAAGCCGAATCTCTCGAATTTG 598			
Db	30 GCAAGACCTTCGACATATCGCGCGCCCACTGATACCCCGAACCTGTCGCAATTCG 89			
Qy	599 AAGCTGTTGCGGTAATATGTAAGACCGAAGAGATGTTGAGCGCGCATGAACCTGA 658			
Db	90 AGACATATGTCGCGCTTGGCGCCGACGAGCCGAATCTGTCGCAAGGCGACGCGCTGA 145			
Qy	659 TTGCGGATTAAGAACTCTCGCTCTCTGTTAGTGAACCGCTTCGAAACAGGATATGTCGTCG 718			
Db	150 TGAACGAATCGACTCGCGGCGCTTGTGTGTGAACCGCGGAGACATGACATGACCCGCG 209			
Qy	719 TGCAACCGGCTAAAGCGCCGCTGATATGCCAACCCAGCGCAGAGAAATGATGACGTTA 778			
Db	210 TCCTCCATATGCGACGCGCGCTGACATGCGCGCGCGCGGAGAGTGTTCGACATCA 269			
Qy	779 CCGGTGGGGCGGACACAGTGAATTTGGCGCTCTCTGCGCGGACAGCTGGCAGCGGTAAATTCG 838			
Db	270 CCGGTGTGCGGATACGATATCTCAACCTCTGCGCGCGCGCTTGGCGCGCGCAGAGAC 329			
Qy	839 TGAAGAAGCTGCTTCTTTGCAATCGCGCGCGCTGATGATGATGATGATGATGATGATGAT 898			
Db	330 TGCCTGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 389			
Qy	899 CTTCCACGGTGTTCGCGCATGAGCTGGAATAATGCTGTAC 937			
Db	390 CCGCGCGATACGAGCGCGCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 428			
RESULT 12				
LOCUS	BZ332786	505 bp	mRNA	linear
DEFINITION	U534d03.x1 Perkins LRH Mus musculus cDNA clone IMAGE:3216101 3'			
	similar to SW:RAE_ECOLI P7658 ADP-HEPTOSE SYNTHASE ;, mRNA			
ACCESSION	BZ332786			

VERSION BE332786.1 GI:9206562  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.  
 1 (bases 1 to 505)  
 REFERENCE Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Strept, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gidbons, M., Page, D., Harvey, N., Schurk, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 The Mashu-NCI Mouse EST Project 1999  
 Unpublished (1999)  
 TITLE Contact: Marra M/Mashu-NCI Mouse EST Project 1999  
 JOURNAL Washington University School of Medicine  
 COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LNL, contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:1066249  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -40RP from Gibco  
 High quality sequence atp: 338.  
 Location/Qualifiers  
 1..505  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3216101"  
 /sex="female"  
 /tissue\_type="primary sorted bone marrow cells"  
 /dev\_stage="adult"  
 /lab\_host="DHI0B (phage-resistant)"  
 /clone\_lib="Perkins LRH"  
 /note="Vector: pZL1, Site\_1: SalI, Site\_2: EagI; cDNA made by oligo-dT priming. Library amplified by stretch PCR. Subtraction method: Bonaldo, et al., Genome Research 6:791. Library constructed by Dr. Archibald Perkins (Yale University)."  
 ORIGIN  
 Query Match 7.0%; Score 100.8; DB 2; Length 505;  
 Best Local Similarity 61.9%; Pred. No. 1.7e-17;  
 Matches 159; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
 1 ATGAAGTACGCTGCGCAGATTGAACTGACAGAGATGATGCTGATGATGATG 60  
 248 ATGAAGTTCATGATCCCGCATGATCAAGCCCTGCTGCTGCTGCTGCTGCTG 307  
 61 CTGATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTA 120  
 308 CTCGACGCTTATCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 367  
 121 GTTAAAGTAACTACATGAAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 368 GTCAATGTCAGCAATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427  
 181 GCTTCTCTGCGATGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 428 GNCCTCTGGGGGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487  
 241 GCGCTGACTAATCTCT 257  
 488 AGCGTACCAACAGTCT 504  
 RESULT 13  
 B2648805

LOCUS B2648805 226 bp DNA linear GSS 29-JAN-2003  
 DEFINITION OGCAM49TM ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0130J01,  
 genomic survey sequence.  
 ACCESSION B2648805  
 VERSION B2648805.1 GI:28112865  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 226)  
 REFERENCE White, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 TITLE Other GSSs: OGCAM49TC  
 JOURNAL Contact: Cathy White  
 COMMENT TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: white@tigr.org  
 Seq primer: TR  
 Class: methylation filtered.  
 Location/Qualifiers  
 1..226  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0130J01"  
 /clone\_lib="ZM 0.7.1.5 KB"  
 /note="Vector: pBCK-; Site\_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"  
 ORIGIN  
 Query Match 6.7%; Score 96; DB 9; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-16;  
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1123 AACAGCAGTCTCCACCAACGCTGAAGGAGATTCCCGGTAACCACTCGAA 1182  
 131 AACAGCAGTCTCCACCAACGCTGAAGGAGATTCCCGGTAACCACTCGAA 190  
 1183 CAGCGTATGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218  
 191 CAGCGTATGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 226  
 RESULT 14  
 LOCUS B2648800/c 251 bp DNA linear GSS 29-JAN-2003  
 DEFINITION OGCAM49TC ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0130J01,  
 genomic survey sequence.  
 ACCESSION B2648800  
 VERSION B2648800.1 GI:28112874  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 251)  
 REFERENCE White, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 TITLE Other GSSs: OGCAM49TM  
 JOURNAL Contact: Cathy White  
 COMMENT TIGR



**THIS PAGE LEFT BLANK**





Db 241 GCGCTGAGTAAATCTCTGGCCGACGTCAAGTCAAAATGCCGCTTGTTCTGTACCGACG 300  
Qy 301 CATCCGACCATTAACCAATTAACGGGTACTTTCCCGCAACCAACGCTGATCCGCTGAT 360  
Db 301 CATCCGACCATTAACCAATTAACGGGTACTTTCCCGCAACCAACGCTGATCCGCTGAT 360  
Qy 361 TTGGAAGAAGGTTTGAAGGTGTGATCCGACGCGCTGCAGACGAGATTAATCAGGCG 420  
Db 361 TTGGAAGAAGGTTTGAAGGTGTGATCCGACGCGCTGCAGACGAGATTAATCAGGCG 420  
Qy 421 CTGAGTTGATTTGCGCGCTGTGCTTTCTGACTACGCCAAAGGTGCGCTGCAAGCGTA 480  
Db 421 CTGAGTTGATTTGCGCGCTGTGCTTTCTGACTACGCCAAAGGTGCGCTGCAAGCGTA 480  
Qy 481 CAGAGATGATCCAACTGGCGCGTAAAGCGGTTTCCGCTGATGATCCAAAGGT 540  
Db 481 CAGAGATGATCCAACTGGCGCGTAAAGCGGTTTCCGCTGATGATCCAAAGGT 540  
Qy 541 ACCGATTTTGAAGCGCTACCGCGCGCTACCGCTTTAAACCGCAATCTCTCGAAATTTGAA 600  
Db 541 ACCGATTTTGAAGCGCTACCGCGCGCTACCGCTTTAAACCGCAATCTCTCGAAATTTGAA 600  
Qy 601 GCTGTGTGCTGATTAATGTAGACCGAAGAGATTTGAGCGCGCGATGAACTGATT 660  
Db 601 GCTGTGTGCTGATTAATGTAGACCGAAGAGATTTGAGCGCGCGATGAACTGATT 660  
Qy 661 GCGCATTTACGAACCTCTGGGCTCTGTGTAGTACCGGTTCCGAACGAGGTAATGTCGCTG 720  
Db 661 GCGCATTTACGAACCTCTGGGCTCTGTGTAGTACCGGTTCCGAACGAGGTAATGTCGCTG 720  
Qy 721 CAACCGGGTAAAGCGCGCTGCATATGCAACCAACGCGCAGAGATGTATGACGTTACC 780  
Db 721 CAACCGGGTAAAGCGCGCTGCATATGCAACCAACGCGCAGAGATGTATGACGTTACC 780  
Qy 781 GGTGCGGGCGACACGCTGATTTGGCGTCTGTGCGGCAACGCTGCGAGCGGTAATTCCTG 840  
Db 781 GGTGCGGGCGACACGCTGATTTGGCGTCTGTGCGGCAACGCTGCGAGCGGTAATTCCTG 840  
Qy 841 GAAAGAGCTGCTCTTTGCCAATGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 841 GAAAGAGCTGCTCTTTGCCAATGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Qy 901 TCCACGGTTTGCAGCATGAGCTGGAATAATGCTGTACGTGACGCTGACATACAGCTTT 960  
Db 901 TCCACGGTTTGCAGCATGAGCTGGAATAATGCTGTACGTGACGCTGACATACAGCTTT 960  
Qy 961 GCGCTGATGACCGAAGAGAACTGAAAGTGGCCGTAGCGGCAAGCGCTTAAAGTGTGAA 1020  
Db 961 GCGCTGATGACCGAAGAGAACTGAAAGTGGCCGTAGCGGCAAGCGCTTAAAGTGTGAA 1020  
Qy 1021 AAAGTGTGATGACCAACGCTGCTTTGACATCCGACGCGCGGCAAGCTCTTATCTG 1080  
Db 1021 AAAGTGTGATGACCAACGCTGCTTTGACATCCGACGCGCGGCAAGCTCTTATCTG 1080  
Qy 1081 GCAAAATGCGCGAAGCTGGTGAACGCTTGAATTTGCGCTCAACAGCGATGCTTCAACC 1140  
Db 1081 GCAAAATGCGCGAAGCTGGTGAACGCTTGAATTTGCGCTCAACAGCGATGCTTCAACC 1140  
Qy 1141 AAAGCGCTGAAAGGGGATTTCCCGCCCGGTAAACCACTCGAACAGCGTATGATGTGCTG 1200  
Db 1141 AAAGCGCTGAAAGGGGATTTCCCGCCCGGTAAACCACTCGAACAGCGTATGATGTGCTG 1200  
Qy 1201 GCGGCACTGGAAGGGGTGACCTGGGTATGTCGTTTGAAGAGGACAGCGCGCAAGCGCTG 1260  
Db 1201 GCGGCACTGGAAGGGGTGACCTGGGTATGTCGTTTGAAGAGGACAGCGCGCAAGCGCTG 1260  
Qy 1261 ATCCCGGGATCTTTCAGATCTGCTGTGAAAGCGCGGACTTAAACGAAAGATTT 1320  
Db 1261 ATCCCGGGATCTTTCAGATCTGCTGTGAAAGCGCGGACTTAAACGAAAGATTT 1320  
Qy 1321 GCGGGAAGTAAAGAGTCTGGGCGCAACGCTGCGAAGTGTGTGCTCACTTTGAAGAC 1380  
Db 1321 GCGGGAAGTAAAGAGTCTGGGCGCAACGCTGCGAAGTGTGTGCTCACTTTGAAGAC 1380

Db 1321 GCGGGAAGTAAAGAGTCTGGGCGCAACGCTGCGAAGTGTGTGCTCACTTTGAAGAC 1380  
Qy 1381 GGTTCGACGACCAACATCATCAAGAGATCCACAGGATAAAGGCTAA 1434  
Db 1381 GGTTCGACGACCAACATCATCAAGAGATCCACAGGATAAAGGCTAA 1434  
RESULT 2  
US-09-489-039A-2037  
; Sequence 2037, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489, 039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 2037  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-2037  
Query Match 70.6%; Score 1011.8; DB 3; Length 1500;  
Best Local Similarity 81.9%; Pred. No. 5e-285;  
Matches 1166; Conservative 0; Mismatches 257; Indels 0; Gaps 0;  
Qy 1 ATGAAAGTAACTGCTGACAGATTTGAACGTGACGAGATGCTGTGATGATGATG 60  
Db 67 ATGAAAGTAACTGCTGACAGATTTGAACGTGACGAGATGCTGTGATGATGATG 126  
Qy 61 CTGAGTCTGATCTGTGACGCGCCCAACGATCTGCTGCGGAAAGCGCTGCGCTG 120  
Db 127 CTGAGTCTGATCTGTGACGCGCCCAACGATCTGCTGCGGAAAGCGCTGCGCTG 186  
Qy 121 GTTAAAGTAACTGCTGACAGATTTGAACGTGACGAGATGCTGTGATGATGATG 180  
Db 187 GTTAAAGTAACTGCTGACAGATTTGAACGTGACGAGATGCTGTGATGATGATG 246  
Qy 181 GCTTCTCTGCTGCTGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
Db 247 GCTTCTCTGCTGCTGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
Qy 241 GCGCTGATTAATCTCTGCGCGACGCTCAACGTAATGCAATCTTCTGTTCTGTAACG 300  
Db 307 GCGCTGATTAATCTCTGCGCGACGCTCAACGTAATGCAATCTTCTGTTCTGTAACG 366  
Qy 301 CATCCGACCATTAACCAATTAACGGGTACTTTCCCGCAACCAACGCTGATCCGCTGAT 360  
Db 367 CATCCGACCATTAACCAATTAACGGGTACTTTCCCGCAACCAACGCTGATCCGCTGAT 426  
Qy 361 TTGGAAGAAGGTTTGAAGGTGTGATCCGACGCGCTGCAGACGAGATTAATCAGGCG 420  
Db 427 TTGGAAGAAGGTTTGAAGGTGTGATCCGACGCGCTGCAGACGAGATTAATCAGGCG 486  
Qy 421 CTGAGTTGATTTGCGCGCTGTGCTTTCTGACTACGCCAAAGGTGCGCTGCAAGCGTA 480  
Db 487 CTGAGTTGATTTGCGCGCTGTGCTTTCTGACTACGCCAAAGGTGCGCTGCAAGCGTA 546  
Qy 481 CAGAGATGATCCAACTGGCGCGTAAAGCGGTTTCCGCTGATGATCCAAAGGT 540  
Db 547 CAGAGATGATCCAACTGGCGCGTAAAGCGGTTTCCGCTGATGATCCAAAGGT 606  
Qy 541 ACCGATTTTGAAGCGCTACCGCGCGCTACCGCTTTAAACCGCAATCTCTCGAAATTTGAA 600  
Db 607 ACCGATTTTGAAGCGCTACCGCGCGCTACCGCTTTAAACCGCAATCTCTCGAAATTTGAA 666  
Qy 601 GCTGTGTGCTGATTAATGTAGACCGAAGAGATTTGAGCGCGGATGAACTGATT 660  
Db 601 GCTGTGTGCTGATTAATGTAGACCGAAGAGATTTGAGCGCGGATGAACTGATT 660

Db 667 GCGGTGTGGGCAAGTCCAGATGAAGCGCAATGCTGAGCGCGCATGAAGCTCATC 726  
Qy 661 GCGGATTACGAATCTTGGCTCTGTTAGTGAACCGCTTCCGAACAGGGTATGTGCTGCTG 720  
Db 727 GCGGAATTCACATCTGTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786  
Qy 721 CAACCGGGTAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 787 CAGCGGAGACGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846  
Qy 781 GGTGCGGGCGACACGTGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
Db 847 GCGCGCGCGACACGTGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906  
Qy 841 GAAGAAAGCTGCTTCTTGGCAATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 907 GAAGAGGCTGCTGATTTGCTGAACGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 966  
Qy 901 TCACCGGTTTCCCGATCGAGCTGGAATAATGCTGTAAGTGAAGTGAAGTGAAGTGAAG 960  
Db 967 TCACCGGTTTCCCGATCGAGCTGGAATAATGCTGTAAGTGAAGTGAAGTGAAGTGAAG 1026  
Qy 961 GCGGTGATGACCGAAGAGAACTGAAGCTGCGCTGACCGCGACGCGCTGAACGCTGTGA 1020  
Db 1027 GCGGTGATGACCGAAGAGAACTGAAGCTGCGCTGACCGCGACGCGCTGAACGCTGTGA 1086  
Qy 1021 AAAGGTGATGACCGAAGAGAACTGAAGCTGCGCTGACCGCGACGCGCTGAACGCTGTGA 1080  
Db 1087 AAAGGTGATGACCGAAGAGAACTGAAGCTGCGCTGACCGCGACGCGCTGAACGCTGTGA 1146  
Qy 1081 GCAATGACCGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 1147 GCGAATGCGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206  
Qy 1141 AAACGCTGAAGAGAGATTCGCGCGCGCTGAACCGCACTGCAACGCTGATGATGCTG 1200  
Db 1207 AAACGCTGAAGAGAGATTCGCGCGCGCTGAACCGCACTGCAACGCTGATGATGCTG 1266  
Qy 1201 GCGGCACTGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
Db 1267 GCGGCGCTGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326  
Qy 1261 ATGCGCGGAGCTTTCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
Db 1327 ATGCGCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386  
Qy 1321 GCGCGGAGTAAAGAGTCTGCGCGCAACGCTGCGCAAGTGTGTGCTCACTTTGAAGAC 1380  
Db 1387 GCGCGGAGTAAAGAGTCTGCGCGCAACGCTGCGCAAGTGTGTGCTCACTTTGAAGAT 1446  
Qy 1381 GGTGCTGCAACCAATCATCAAGAGATCCACAGAGATA 1423  
Db 1447 GGTGCTGCAACCAATCATCAAGAGATCCACAGAGATA 1489

RESULT 3  
US-09-489-039A-2107/c  
Sequence 2107, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
PRIOR APPLICATION NUMBER: 2000-01-27  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 2107  
LENGTH: 1323  
TYPE: DNA  
ORGANISM: Klebsiella pneumoniae

US-09-489-039A-2107  
Query Match 60.7%; Score 870.6; DB 3; Length 1323;  
Best Local Similarity 81.8%; Pred. No. 8.1e-244;  
Matches 1005; Conservative 0; Mismatches 224; Indels 0; Gaps 0;  
Qy 1 ATGAAGTAAAGCTGCGCAAGGTTGAACGTGACAGAGTATGCTGCTGCTGCTGCTGCTG 60  
Db 1230 ATGAAGTAAAGCTGCGCAAGGTTGAACGTGACAGAGTATGCTGCTGCTGCTGCTGCTG 1171  
Qy 61 CTGAATCTTACTGTAGAGCGCCCAACAGTGTATCTGCGCGAAGCGCGCTGCTGCTG 120  
Db 1170 CTGAATCTTACTGTAGAGCGCCCAACAGTGTATCTGCGCGAAGCGCGCTGCTGCTG 1111  
Qy 121 GTTAAAGTAAATCCATTCGAAGAACTCCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 180  
Db 1110 GTTAAAGTAAATCCATTCGAAGAACTCCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 1051  
Qy 181 GCTTCTCTGCTGCTATGACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
Db 1050 GCTTCTCTGCTGCTATGACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 991  
Qy 241 GCGCTGATGAATCTGCGCGCAAGCTGAACGCTGAATGCGAATGCGAATGCGAATGCGA 300  
Db 990 GCGCTGATGAATCTGCGCGCAAGCTGAACGCTGAATGCGAATGCGAATGCGAATGCGA 931  
Qy 301 CATCCGACATTAACCAATTAAGGCTATCTTCCCGCAACCAACGCTGATCCGCTGCTGAT 360  
Db 930 CATCCGACATTAACCAATTAAGGCTATCTTCCCGCAACCAACGCTGATCCGCTGCTGAT 871  
Qy 361 TTTGAAGAAAGTTCGAAGGCTGATGCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
Db 870 TTTGAAGAAAGTTCGAAGGCTGATGCTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811  
Qy 421 CTGAGTTCATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 810 CTGAGTTCATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751  
Qy 481 CAGCAATGATCCAACTGCGCGCTGAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db 750 CAGCAATGATCCAACTGCGCGCTGAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 691  
Qy 541 ACCGATTTTGAAGGCTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 690 ACCGATTTTGAAGGCTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631  
Qy 601 GCTGTTGCTGTAATGTAAGACCGAAGAAAGATTTGAGCGCGCATGAAGCTGAT 660  
Db 630 GCTGTTGCTGTAATGTAAGACCGAAGAAAGATTTGAGCGCGCATGAAGCTGAT 571  
Qy 661 GCGGATTAGAACTCTGCGCTCTGTTAGTGAACCGTTCCGAACAGGCTATGTGCTGCTG 720  
Db 570 GCGGATTAGAACTCTGCGCTCTGTTAGTGAACCGTTCCGAACAGGCTATGTGCTGCTG 511  
Qy 721 CAACCGGGTAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 510 CAACCGGGTAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451  
Qy 781 GGTGCGGGCGACACGTGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
Db 450 GGTGCGGGCGACACGTGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391  
Qy 841 GAAGAAAGCTGCTTCTTGGCAATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 390 GAAGAAAGCTGCTTCTTGGCAATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 331  
Qy 901 TCACCGGTTTCCCGATCGAGCTGGAATAATGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGA 960  
Db 330 TCACCGGTTTCCCGATCGAGCTGGAATAATGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGA 271  
Qy 961 GCGGTGATGACCGAAGAGAACTGAAGCTGCGCTGAGCGCGACGCGCTGAACGCTGTGA 1020  
Db 270 GCGGTGATGACCGAAGAGAACTGAAGCTGCGCTGAGCGCGACGCGCTGAACGCTGTGA 211

Qy 1021 AAAGTGTGATGACCAACGAGTGTCTTGAACATCTGACACGCGGACAGTCTTATCTG 1080  
Db 210 AAAGTGTGATGACCAACGAGTGTCTTGAACATCTGACACGCGTCAATGTCTCTATCTG 151  
Qy 1081 GCAAAATCCCGCAAGCTGGGTGACCGCTTGAATGTTGTCCTGCAACAGCAGATGCTTCAAC 1140  
Db 150 GCGAATGCGGCAAGCTGGGTGACCGCTTGAATGTTGTCCTGCAACAGCAGATGCTTCAAC 91  
Qy 1141 AAAGCGTGAAGGGGAAATCCCGCGGTAAACCACTCGAAGCGGTATGATGATGCTG 1200  
Db 90 AAAGCTGAAGGGGAAATCCCGCGGTAAACCACTCGAAGCGGTATGATGATGCTG 31  
Qy 1201 GCGGCACTGGAAGCGGTGACCTGAGTGTAGT 1229  
Db 30 GCGGCGCTGGAAGCGGTGACCTGAGTGTAGT 2

RESULT 4  
US-09-543-681A-2462  
; Sequence 2462, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 2462  
; LENGTH: 1476  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
; US-09-543-681A-2462

Query Match 50.5%; Score 724.2; DB 3; Length 1476;  
Best Local Similarity 69.3%; Pred. No. 4,76-201;  
Matches 987; Conservative 0; Mismatches 438; Indels 0; Gaps 0;

Qy 1 ATGAAAGTAAAGCTGCGAAGTGTGAACGTGACAGATGATGATGATGATGATGATG 60  
Db 52 ATGAAAGTAAAGCTGCGAAGTGTGAACGTGATGATGATGATGATGATGATGATG 111  
Qy 61 CTGATGCTTACTGCTGACGCGCCCAACGATGATGATGATGATGATGATGATGATG 120  
Db 112 TTAGACCGTATATGCTGATGCTTCAAGATGATGATGATGATGATGATGATGATG 171  
Qy 121 GTTAAAGTAAATCATGAAAGAGTCCCGGCGCGGCTTAACGTGCGATGATATATC 180  
Db 172 GTTAAAGTAAATCATGAAAGAGTCCCGGCGCGGCTTAACGTGCGATGATATATC 231  
Qy 181 GCTTCTCTGCTGATGATGACGCGCTGCTGCTGATGATGATGATGATGATGATG 240  
Db 232 GCTTCTTAAAGGCGCTAAATCTGTTGGTGGGTTAACCGGATGATGATGATGATG 291  
Qy 241 GCGGTGATTAATCTCTGCGCGAGTCAACGTCAATGCACTTCTGATGATGATG 300  
Db 292 GCAATTAAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 351  
Qy 301 CATCGACCATTAATCAATTAATGAGGATCTTCCGCAACCAAGCTGATCTGATGAT 360  
Db 352 CATCGACCATTAATCAATTAATGAGGATCTTCCGCAACCAAGCTGATCTGATGAT 411  
Qy 361 TTGTAAGAAAGTTTGAAGGTTGATGATGATGATGATGATGATGATGATGATGATG 420  
Db 412 TTGTAAGAAAGGATTTAGCAATGCTCTCCAGAGCAATTAATGAGGATTAATCAAGCA 471  
Qy 421 CTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
Db 472 TTACCTCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531

Qy 481 CAGCAGATGATCCAACTGCGCGGTAAAGCGGTGTTCCGGTGTGATGATGATGATGATG 540  
Db 532 CAGCAGATGATCCAACTGCGCGGTAAAGCGGTGTTCCGGTGTGATGATGATGATGATG 591  
Qy 541 ACCGATTTTGAAGCGCTTACCGCGCGCTTACCGCTTAAACCGCATCTCTCGAATTTGA 600  
Db 592 AGTGAATTTGAAGCGCTTACCGCGCGCTTACCGCTTAAACCGCATCTCTCGAATTTGA 651  
Qy 601 GCTGTTGCGGTAAATGTAAGACCGAAGAGATGTTGAGCGCGCATGAAATCGATG 660  
Db 652 CAAGTGTGCGGTGTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATG 711  
Qy 661 GCGATTTAAGCACTCTGCGCTCTGATGATGATGATGATGATGATGATGATGATGATG 720  
Db 712 AAAGATTTGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771  
Qy 721 CAACCGGATTAAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 780  
Db 772 CAAGCAATGAAGCGCGCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 831  
Qy 781 GGTGCGGCGGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 832 GGGCGCGGAGATGCGGTATGCGCGCTTCTGATGATGATGATGATGATGATGATGATG 891  
Qy 841 GAAGAGCTGCTTCTTGGCAATGCGCGCGCTGCGGTGCTGCGCAACTGGGAAC 900  
Db 892 GCGGAAGCGTCTTCTTGGCAATGCGCGCGCTGCGGTGCTGCGCAACTGGGAAC 951  
Qy 901 TCGACGTTTCCGCGCATGCGATGGAATAATGCTGTATGATGATGATGATGATGATG 960  
Db 952 TCGACGTTTCTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1011  
Qy 961 GCGGTATGACCGAAGAGAACTGAAGCTGCGCGGTGCGGCAAGCGGTAAACGTGTGAA 1020  
Db 1012 GGTATTAATGAGAGATGCAACATGCAACCAAGCGGTGCTTGAAGCGGTGATGATG 1071  
Qy 1072 CGCATGATGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1131  
Qy 1081 GCAAAATCCCGCAAGCTGGGTGACCGCTGATGATGATGATGATGATGATGATGATG 1140  
Db 1132 GCTAAATGCAAGTAAATGCGGTGATGCGCTGATGATGATGATGATGATGATGATG 1191  
Qy 1141 AAAGCGTGAAGGGGATCCCGCGGTAAACCACTGCAACAGGATGATGATGATG 1200  
Db 1192 AAAGCATTAAGGTGAAGTCCCGCATGCAACCTTGAAGCAAGATGACGATTTA 1251  
Qy 1201 GCGGCACTGGAAGCGGTGCACTGCGGTATGCTGTTGAAGAGCAACGCGGCGCTTG 1260  
Db 1252 GCGCGCTTGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1311  
Qy 1261 ATGCGCGGATCTTGGCAATGCTGCTGTGAAGCGCGGATGATTAACGAAGAGAT 1320  
Db 1312 ATGCAATCAATATGCTGATATTTTAAGAGGTGCGGATTAATTAACCGAAGCAAT 1371  
Qy 1321 GCGGAGATTAAGAGTCCGCGCAACGATGCGGAAGTGTGCTGCACTTTGAAGAC 1380  
Db 1372 GCTGTGATTAAGAGTGTGCGGCTGCGGATGATGATGATGATGATGATGATGATG 1431  
Qy 1381 GGTGCTGACGACCAATCATCAAGAAATCAACAGATTA 1425  
Db 1432 GGTATTTCCACATTAATTAATTAATGCAATTAAGAAATTA 1476

RESULT 5  
US-09-557-884-1/c  
; Sequence 1, Application US/09557884  
; Patent No. 650581  
; GENERAL INFORMATION:  
; APPLICANT: Pletschmann et al.  
; TITLE OF INVENTION: The Nucleotide sequence of



```

Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
Query Match 44.3%; Score 635.2; DB 3; Length 1830121;
Best Local Similarity 65.9%; Pred. No. 1e-173;
Matches 922; Conservative 0; Mismatches 478; Indels 0; Gaps 0;
QY 17 CAGAGTTGAAAGTGCAGAGTGAATGTTGTTGATGTGATCTGATCGTTACTGCT 76
Db 1596554 CAGATTAAAGCAAGCAAAAGTACTCGTATTAGCGAGTGTATGATCGTTATTGGT 1596495
QY 77 ACCGCCCCACAGTCTGATCTCGCCGGAAGCGCGTCCCGTGTAAAGTGAATACCA 136
Db 1596494 TCGGCGCAACCAACCGTATTTCACAGAAAGCACAGTACCGTGTGTGTACAGAAA 1596435
QY 137 TCGAAGAACTCGCGGCGCGCGCTAACTGCGATGATATGCTTCTCGGTGCTA 196
Db 1596434 ATGAAGAAACGCGCGGTGTGTGACAGCAAAATGCGAGTGAATATGCTTCACTCATATGAC 1596375
QY 197 ATGACCGCTGTGTGCGGTGTAACGAGCATTAAGTACGACGCGCGCTGAATATCTC 256
Db 1596374 CCGTTCAGTTAATGGGATTTGATTGGAACAAATGAACTGTGCACTTCCCTTTAT 1596315
QY 257 TGGCGCAGTCAAGTCAAAATGCACTTCGTTTCTGTATCCGACGATCCGACATTAACA 316
Db 1596314 TAGAAAAACAAAATGATTGTAATTTTGTGATTAAGAAACCATCAACATTAATA 1596255
QY 317 AATTAACGGTACTTTCGCAACCAACAGTATCCGCTGATTTTGAAGAAGTTTGG 376
Db 1596254 AATTACGATTATTCTGTGATCAACAGCTGCTCCGCTTGATTTTGAAGAAGATTTC 1596195
* QY 377 AAGGTGTTGATCCGACCGCTGCAAGACGAGTTAATCAAGGCGCTGAATTTGATGGCG 436

```

```

Db 1596194 ATAATGATGATGCAAGATTTAATTAGCAAGTTAGAAAAGTGGCGTGAATAAATTAACGTTG 1596135
QY 437 CGGTGGTCTTTTGTGCTACGCCAAGGTGGCGTGGCAAGCGTACAGCATGATCCAAC 496
Db 1596134 CTTTGATTTCTTTGTGATTTAGCGCAAGAGCAAGCTTTAAAGTTTTCAGAAATATTCAAA 1596075
QY 497 TGGCGGTAAAGCGGGTGTCCGCTGTGATGATCCAAAAGGTACCGATTTTGAACGCT 556
Db 1596074 TTGCAGCGAAAGGAATGTGCTGTGTGATGATCCAAAAGGGAATGATTTGAACGTT 1596015
QY 557 ACCGCGCGTACCGCTGTTAACCGCGAATCTCTCGGAATTTGAAGCTGTGTGTGAAT 616
Db 1596014 ATCGTGGGCTACATTATTGACACCCAAATATGCTCAATTTGAAGCTGTGTGTGAAT 1595955
QY 617 GTAAAGCCGAAGAAAGATTGTTGAGCGGCGCATGAACCTGATGCGAATTAACAATCT 676
Db 1595954 GCAATACGGAAGAAAGATTTATTGAAAGGTTTAAATTAATTTCTGATATTGAATTA 1595895
QY 677 CGGCTGTGTTAGTGAACCGCTTCCGAACAGGTATGTGCTGCTGCAACCGGCTAAAGCG 736
Db 1595894 CGCACATTTTGTGAGCGGCTTGAATAAGGATGACATTATTAACGCCCAATCAAGAGC 1595835
QY 737 CGCTGCATATGCCAACCCCAAGCGCAGAAAGTATGACGTTACCGGTGGCGGCGACCG 796
Db 1595834 CTTATCATTTGCGCAACGTTGCAAAAAGAGTGTGATGACGAGGACCGGTGACACTG 1595775
QY 797 TGATGGCCGCTCGGCGGCAACGCTGCAAGCGGGTAAATTCGTGAAGAAAGCTGCTTCT 856
Db 1595774 TCAATTACGATTATTAGCAACCGCATTAAGCATGAGAGTCTTTGCAAGAACTTTGTTACC 1595715
QY 857 TTGCAATGCGGCGGCTGGCGGTGTGTGTCGCAAACTGAGAACTTCACGTTTCCGCA 916
Db 1595714 TAGCAATGTGTGCGCAGGAATGTGTGGGGAATTTGGGGAATTCACAGCGTTTCAACG 1595655
QY 917 TCGAGCTGAAAAATGCTGTACGTGACGTGCAATACAGCTTTGGCGGTATGACCGAAG 976
Db 1595654 TGAACCTTGAAATGGAATTCAGTCTGCTGTAACCTGATTTGGCAATTATAGTGAAG 1595595
QY 977 AGGACGTGAAGCTGGCGGTAGCGGCGGCGCGGTAAAGTGTGTAAGAAAGTGTATGACCA 1036
Db 1595594 CAGAAATTAAGAAATGCTGTGCAACAGCTTAAGCGGCGGTGAAAAATTTGTATGACTA 1595535
QY 1037 ACGGTGCTTTGACATCTGCAACGCGGCAAGCTCTTATCTGGCAAAATGCGCAAGC 1096
Db 1595534 ATGGCTGTTTGTATATTGTCATCAAGGCGATTTTCTTATTAAGAAAATGACGCAAT 1595475
QY 1097 TGGGTGACGCTTGATTTGTTGCCGTCAACAGCGATGCTTCAACAAAACGCTGAAGGGG 1156
Db 1595474 TGGGCGATCTCTAATTTGTCGCTAAACAGCAGCAATCTCTGTTAAACGCTTAAAGGTG 1595415
QY 1157 ATTCGCGCCGCGTAACCCCACTGAAACAGGCTATGATTTGTCGCGCACTGAAGCGG 1216
Db 1595414 AAGTGCGCCCAATTAATTAATCTTGAACCGTATGCGGCTATTTGCTGTGTTGACATCG 1595355
QY 1217 TCGATGCGGTAGTGTGTTGAAGAGACACGCGCAGCGCTTGAATCGCGGATCTTGC 1276
Db 1595354 TAGATGCTGTGTGTCCTTCACTGAAGATACCAACAAAGCTTAAATCGGGAATTTTAC 1595295
QY 1277 CAGATCTGCTGTGAAGAGGCGGCACTATTAACAGAAAGATTTGCCGGAATTAAGAG 1336
Db 1595294 CAGATCTTTTATGCAAAAGGCGGATTAACAAACGAAAGATTTGACGGGAGTAAAG 1595235
QY 1337 TCTGGGCAACGCTGGCGAAGTGTGCTCAACTTTGAAGAAGCTGTGCTGACGACCA 1396
Db 1595234 TTTGGGCAACGCTGGCGATGTTAAAGTGTCTAACTTTGAAGATGTTGTTCAACAA 1595175
QY 1397 ACATCATCAAGAGATCCA 1416
Db 1595174 ATGTAATGAAAAAATTA 1595155

```

RESULT 7



LOCATION:	(45732)..(45732)
OTHER INFORMATION:	n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(47036)..(47036) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(51334)..(51334) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(51602)..(51602) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(51805)..(51805) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(55369)..(55369) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(65309)..(65309) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(65313)..(65313) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(80024)..(80024) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(100091)..(100091) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(102696)..(102696) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(105121)..(105121) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(107248)..(107248) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(117136)..(117136) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(119750)..(119750) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(119524)..(119524) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(120038)..(120038) OTHER INFORMATION:n equals a,t,c, or g
FEATURE:	
NAME/KEY:	misc.feature LOCATION:(121344)..(121344)



Db 1595294 CAGATCTTTAGTGAAGGGGGGATTAACAACCCGAAAGATTGACGAGTAAGAAG 1595235  
 Qy 1337 TCTGGCCAAACGGTGGGAAAGTGTGCTCAACTTTGAAGAGGTTGCTCGACGACCA 1396  
 Db 1595234 TTTGGGCAAAACGGTGGGAAAGTGTGCTCAACTTTGAAGAAATGTTTCAACAA 1595175  
 Qy 1397 ACATCATCAAGAGATCCAA 1416  
 Db 1595174 ATGTGATTGAAAATAATTTAA 1595155

## RESULT 8

US-09-252-991A-16413  
 / Sequence 16413, Application US/09252991A  
 / Patent No. 6551795  
 / GENERAL INFORMATION:  
 / APPLICANT: Marc J. Rubenfield et al.  
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 / TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 / FILE REFERENCE: 107196.136  
 / CURRENT APPLICATION NUMBER: US/09/252,991A  
 / CURRENT FILING DATE: 1999-02-18  
 / PRIOR APPLICATION NUMBER: US 60/074,788  
 / PRIOR FILING DATE: 1998-02-18  
 / PRIOR APPLICATION NUMBER: US 60/094,190  
 / PRIOR FILING DATE: 1998-07-27  
 / NUMBER OF SEQ ID NOS: 33142  
 / SEQ ID NO 16413  
 / LENGTH: 1581  
 / TYPE: DNA  
 / ORGANISM: Pseudomonas aeruginosa  
 / US-09-252-991A-16413

Query Match 34.5%; Score 494.2; DB 3; Length 1581;  
 Best Local Similarity 59.8%; Pred. No. 6.9e-134;  
 Matches 848; Conservative 0; Mismatches 568; Indels 3; Gaps 1;

Qy 1 ATGAAGTAAACGCTGCCAGATTGAACTGTCAGAGATGATGCTGCTGATGATGATG 60  
 Db 160 ATGAAGTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219  
 Qy 61 CTGATGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120  
 Db 220 CTGATGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 219  
 Qy 121 GTTAAAGTAAATCATGCAAGAACTGCGGCGCGCGCGCTTAACTGCGGATGATATC 180  
 Db 280 GTCCGCGTCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAATC 339  
 Qy 181 GCTTCTCTCGGTGTAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Db 340 GCGCGCTGCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 399  
 Qy 241 GCGCTGATTAATCTGCGCGCAAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 300  
 Db 400 AGCTGCGCAACAGCTCAAGGCGCGCTGAGTGAAGCGCGCTTCAAGCGCATGATAGC 459  
 Qy 301 CATCCAGCATTAACAATTAACGGGTAATTCGCGCAACCAACGCTGATCCGCTGAT 360  
 Db 460 CAGCGCAGCATGCTCAAGCTGCGGCTCATGATGCTCAACGAGCAATGCTGCGGCTGAC 519  
 Qy 361 TTTGAAGAAGTTTCAAGGTTGATCCGCAAGCGCTGCAAGCGGATTAATCAAGCG 420  
 Db 520 TTCGAGGAAACCGTTC--CGACCGAAGCGCGCGCGCTGCGCTGAGAGTCTGCTG 576  
 Qy 421 CTGAGTTCAATGCGCGCTGCTGCTTCTGATCAACCAAGTGCCTGCGCAAGCGTA 480  
 Db 577 CTGCGCAAGTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636  
 Qy 481 CAGCAGATGATCAACTGCGCGCTGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 Db 637 CAGGTGCTGATCCAGCGCGCGCGCGCGCGCAATCTCGGTAATGCGGATCCCAAGGCG 696

Qy 541 ACCGATTTGAGCGCTACCGCGCGCGCTACGCTGTAACCGCAATCTCTCGAATTTGA 600  
 Db 697 AAGACTTGGCCATCTATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 756  
 Qy 601 GCTGTTGCTGCTAAATGTAAGCCGAAGAAAGATTTGAGCGCGCGCATGAATGAT 660  
 Db 757 ACCATGTCGCGCGCTGCG 816  
 Qy 661 GCGGATTAAGATCTTCCGCTGCTGTAAGACCCGCTTCCGAACAGGATATGCTGCTG 720  
 Db 817 AGCGAATCGAATCTTCCGCTGCTGTAAGACCCGCGCGCGCGCGCGCGCGCGCGCG 876  
 Qy 721 CAACCGGTTAAAGCG 780  
 Db 877 GCGATGCGCAAGCG 936  
 Qy 781 GGTGCGGCGCAACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 Db 937 GGTGCGGCGCAACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996  
 Qy 841 GAAGAACTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 Db 997 CCGTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056  
 Qy 901 TCCAGGATTTCCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
 Db 1057 GCGGCGATCAAGCG 1116  
 Qy 961 GCGGATGATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1020  
 Db 1117 GCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176  
 Qy 1021 AAGTGTGATGATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1080  
 Db 1177 AAGTGTGATGATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1236  
 Qy 1081 GCAATGCG 1140  
 Db 1237 GAAAGCG 1296  
 Qy 1141 AAGCGCTGTAAGCG 1200  
 Db 1297 ACTGCGCTGTAAGCG 1356  
 Qy 1201 GCGGCACTGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 Db 1357 GCGGCGCTGCG 1416  
 Qy 1261 ATGCGCGGATCTTTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 Db 1417 CTGAGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1476  
 Qy 1321 GCGGCGGTAAGAGTCTGCGCGCAAGCGTGCAGAGTGTGCTGCTCACTTTGAAGC 1380  
 Db 1477 GTGCGCGCGAGATGCTCAAGGCTCAAGGCGGAGAGTACGCGGTGCTGCGGCTGCTG 1536  
 Qy 1381 GCTTCTGCGAGCAATCATCAAGAAATCAACAG 1419  
 Db 1537 AACAGCTCCACCAACCGCATGCTGCGAAGATCCGCGAG 1575

## RESULT 9

US-09-252-991A-16304  
 / Sequence 16304, Application US/09252991A  
 / Patent No. 6551795  
 / GENERAL INFORMATION:  
 / APPLICANT: Marc J. Rubenfield et al.  
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 / TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 / FILE REFERENCE: 107196.136  
 / CURRENT APPLICATION NUMBER: US/09/252,991A  
 / PRIOR FILING DATE: 1998-02-18  
 / PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 16304  
LENGTH: 1251  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16304

Query Match 27.3%; Score 391; DB 3; Length 1251;  
Best Local Similarity 60.0%; Pred. No. 8.5e-104;  
Matches 671; Conservative 0; Mismatches 445; Indels 3; Gaps 1;

QY 301 CATCGACCATTTACCAATTACGGGTACTTCCGCAACCAAGCTGATCCGCTGGAT 360  
DB 2 CAGCGACATCGTCAAGCTCGGGTCATGAGTGTCAACGACATGCTGCGGGTGCAC 61  
QY 361 TTGAGAGAGGTTTGAAGGTGTGATCCGAGCCGCTGACGAGCGGATTATCAGGCG 420  
DB 62 TTGAGAGAGGTTTGAAGGTGTGATCCGAGCCGCTGAGCTGAGAGTCCGCTG 118  
QY 421 CTGAGTTGATTTGGCGGCTGCTGTTCTGATACGCAAGTGGCTGGCAAGCTA 480  
DB 119 CTGGCAAGGTCAAGGTGCTGTGTCTGTCCGATACGCGCAAGGCGCTACAGAACAC 178  
QY 481 CAGCAGATGATCCAACTGCGCGGTAAAGCGGGTGTCCGCGTGTGATTTGATCCAAAAGT 540  
DB 179 CAGGTGCTGATCCAGCGCGCGCGCGCGCAACATCCGATGTCGCGGATCCAGGGC 238  
QY 541 ACCGATTTGAGCGCTACCGCGCGCTACGCTGTTAACCGCGAAATCTCTCGGAATTTGAA 600  
DB 239 AAGGACTTCGCACTGATCGGGCGCGCGCGCTGATACCGCGAACTGTCCGAATTCGAG 298  
QY 601 GCTGTTGCTGCTAAATGTTAAGACGGAAGAGATTTGTAACCGCGCGCTAAACTGATT 660  
DB 299 ACCATTCGTCGCGCGCTGTCGCGCAAGCGCAACCTGTGTCGCAAGGCGCGCTGATG 358  
QY 661 GCCGATTTACGAATCTCGGCTCTGTTAGTGAACCGGTTCCGAACGGGATGTCGCTG 720  
DB 359 AGCGAATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418  
QY 721 CAACCGGTTAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
DB 419 CGCGATGCGCAAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 478  
QY 781 GGTGCGGCGCAACGCTGATTTGGCGTCTGCGCGCAACGCTGCGAGCGGATTTGCTG 840  
DB 479 GGTGCGGCGCAACGCTGATTTGGCGTCTGCGCGCAACGCTGCGAGCGGATTTGCTG 538  
QY 841 GAAGAAGCTGCTTCTTGAAGTGAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
DB 539 CCGTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598  
QY 901 TCCACGCTTTGCGCGATGAGCTGGAATAATGCTGATCGTGAACGTCGATACAGGCTTT 960  
DB 599 GCGGAGATCAAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658  
QY 961 GCGCTGATGACGGAAGAACTGAAGCTGCGCGCTGAGCGGCAAGCGCTGAAACGCTG 1020  
DB 659 GCGCTGATGACGGAAGAACTGAAGCTGCGCGCTGAGCGGCAAGCGCTGAAACGCTG 718  
QY 1021 AAGTGTGATGACCAACGAGTGTCTTTGACATCTGCAAGCGCGGAGCACTGCTTATCTG 1080  
DB 719 AAGTGTGATGACCAACGAGTGTCTTTGACATCTGCAAGCGCGGAGCACTGCTTATCTG 778  
QY 1081 GCAAAATGCGCGCAAGCTGAGTGAACGCTTGAATGTTGCGCTCAACAGGATGCTTCAAC 1140  
DB 779 GAAAGAGGCGCGCGCAAGCTGAGTGAACGCTTGAATGTTGCGCTCAACAGGATGCTTCAAC 838  
QY 1141 AAAAGGCTGAAAGGAGATTTCCGCGCGGTTAAACCACTGCAACAGCTGATATGTTGCTG 1200  
DB 839 ACTGCGCTGAAAGGAGATTTCCGCGCGGTTAAACCACTGCAACAGCTGATATGTTGCTG 898

QY 1201 GCGCACTGGAAGCGGTGCACTGGGTAGTGTGCTTTGAAGAGACACGCGCAGCGCTTG 1260  
DB 899 GCGGCTGCGCGCGGTGCACTGGGTAGTGTGCTTTGAAGAGACACCTCCGAGCGCTG 958  
QY 1261 ATGCGCGGATCTTGGCAATCTGCTGTGAAAGCGCGGCACTATTAACGAGAGATT 1320  
DB 959 CTCGAGCAGGTGCGTCCGAGCGTGTGTCAAGGCGCGGATTAACGCGCTGAGCAGTGT 1018  
QY 1321 GCGGAGATTAAGATCTGGGCAACGCGTGGCGAAGTGTGTGCTCAACTTGAAGAC 1380  
DB 1019 GTGCGCGCGCAAGTCTCAAGGCTCAAGCGCGGAGTACGCGTGTGCTGCGCTGTGAG 1078  
QY 1381 GATTGCTGACGACCAACATCATCAAGAGATCCACAG 1419  
DB 1079 AACAGCTTCAACACCGCATCTGCAAGAGATCCGCGAC 1117

RESULT 10  
US-09-252-991A-16178/c  
Sequence 16178, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 16178  
LENGTH: 963  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16178

Query Match 24.4%; Score 350.2; DB 3; Length 963;  
Best Local Similarity 60.8%; Pred. No. 6.2e-92;  
Matches 571; Conservative 0; Mismatches 368; Indels 0; Gaps 0;  
QY 481 CAGCAATGATCCAACTGCGCGCTAAAGCGGCTTCCGCTGCTGATTCGAAAGT 540  
DB 951 CAGGTGCTGATCCAGGCGCGCGCGCGCGCAACATTCGCTATGCGGATCCAGAGGC 892  
QY 541 ACCGATTTGAGCGCTACCGCGCGCTACGCTGTTAACCGCGAATCTCTCGAAATTTGAA 600  
DB 891 AAGGACTTCGCACTATGCGCGCGCGCGCGCTGATCAACCGCAACCTGTCCGAATTCGAG 832  
QY 601 GCTGTTGCTGCTAAATGTTAAGACGGAAGAGATTTGTAAGCGCGCGCATGAACTGATT 660  
DB 831 ACCATGCTGCGCGCTGTCGCGCAAGCGCAACCTGATGCGCAAGGCGCGCGCTGATG 772  
QY 772 GCGGATTTACGAATCTCGGCTCTGTTAGTGAACCGGTTCCGAACAGGCTATGCTGCTG 720  
DB 771 AGGGAATGCACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712  
QY 721 CAACCGGTTAAAGCGCGCTGATGATGCAACCGCAAGCGGAGAGTGTATGAGCTTTAC 780  
DB 711 CGGATGCGCAAGCGCGCGCTGCACTGCGCGCGCGGAGGAGTGTTCAGCTCAC 652  
QY 781 GGTGCGGCGCAACGCTGATTTGGCGTCTGCGCGCAACGCTGCAACGCGGTAATTCGCTG 840  
DB 651 GGTGCGGCGCAACGCTGATTTGCACTGCAACGCTGCGCGCGCTGCGCGCGGAGAGCTG 532  
QY 841 GAAGAAGCTGCTTCTTGAAGTGAAGCGCGCTGAGCGTGTGCTGCGCAACTGCGAAGCTG 900  
DB 591 CCGTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 532  
QY 901 TCCACGCTTTGCGCGATGAGCTGGAATAATGCTGATCGTGAACGTCGATACAGGCTTT 960



```

; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15975

```

Query Match	12.3%	Score 176.2;	DB 3;	Length 501;
Best Local Similarity	59.5%	Pred. No. 3e-41;		
Matches 298; Conservative	0;	Mismatches 203;	Indels 0;	Gaps 0;

Qy	555	CTACCGCGGCGCTACGCTGTTAACCCGCAATCTCTCGAATTCTTGAAGCTGTGTGCTGTTAA	614
Db	501	CTATCGCGGCGCGACGCTGATCAACCCCGAACTGTCCGAATTCCAGACCAATCGTGGCCG	442
Qy	615	ATGTAAAGCCGAAGAAGATTGTGAACCGCGGCAATGAACCTGATCCGATTAAGCACT	674
Db	441	TTTGCGCGCAGAAAGCGAACTGTGTCCGAAGGSCCAAGCGCTGATGAAGCGAACTGACCT	382
Qy	675	CTCGGCTCTGTTAAGTACCCCGTTCCGAACAGGGTATGTCTGCTGCTGCACCGGGTAAAGC	734
Db	381	CGGTGCTCTGTGCTGTGAACCCCGCGGAGACATGTGATGACCTGTCTCCGGAAATGGCAAC	322
Qy	735	GCCGCTGCATATGCCAACCCCAAGCGCAGAGAACTGTATGACCTTACCGGTGCGGCGGACAC	794
Db	321	GCGCTGCACCTGCGCGGCGCGGCGCGGAGAACTGTTCACGTCAACCGGTGCGGCGGAAATC	262
Qy	795	GATGATTTGCGGCTGCGGCGGCAAGCTGTGACAGCGGATTAATTGCTGGAAGAAAGCCTGCTT	854
Db	261	GGTCAATCTCAACCTCTGCGCGCGCGCGCTTGTCCGCGCGGACAGACGTGTCCGTCGCGGTGG	202
Qy	855	CTTTGCCAATGCGCGCGCTGTGCGGTGTGTGTGCGCAAACTGGAACTTCAACGCTTTCCGC	914
Db	201	CCTGGCCCAACTGTGCGCGCGCGCATGTGTGTGCGCAAGCTGTGGGTAAACCGCGCGCATACGGC	142
Qy	915	GATCGAGCTGGAATAATGCTGTACGTGTGACGATCAGATACAGGCTTTTGGCGTATGACCGA	974
Db	141	GCCCGAACTTGCGTGGCGCGGCTGCACGCGGAGAGAGGTTTCCAGCTGTGGCGTGTCTGGGCTT	82
Qy	975	AGAGGAACCTGAAGCTGGCCGTGAGCGGACGCGGTAAACGTGTGTGAAAAAGTGTGATGAC	1034
Db	81	GGAGCAATTGCTGTGCTGGCAATCGAAGACGCGCGCCACCGCGAGAAAGATGTTCTTAC	22
Qy	1035	CACCGGTGCTTTGACATCCT	1055
Db	21	CAATGGCTGTTCGACATCCT	1

RESULT 13  
 US-08-651-155B-185  
 Sequence 185, Application US/08651155B  
 Patent No. 6365401  
 GENERAL INFORMATION:  
 APPLICANT: Mahan Dr., Michael J.  
 APPLICANT: Conner Mr., Christopher P.  
 APPLICANT: Hiethoff Mr., Douglas M.  
 TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION  
 TITLE OF INVENTION: OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST  
 TITLE OF INVENTION: INFECTION  
 NUMBER OF SEQUENCES: 255  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Chrisman, Bynum & Johnson, P.C.  
 STREET: 1900 Fifteenth Street  
 CITY: Boulder  
 STATE: CO  
 COUNTRY: USA  
 ZIP: 80302  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/651,155B  
 FILING DATE: 17-MAY-1996  
 CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Peterson Mr., Steven C.  
REGISTRATION NUMBER: 36,238  
REFERENCE/DOCKET NUMBER: 17060.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303/546-1300  
TELEFAX: 303/449-5426  
TELEX: ABH1475  
INFORMATION FOR SEQ ID NO: 185:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULAR TYPE: DNA (genomic)  
HYPOTHEETICAL: NO  
ANTI-SENSE: YES

Query Match	10.4%	Score 149.4	DB 3	Length 347
Best Local Similarity	74.8%	Pred. No. 1.7e-33		
Matches 240	Conservative 0	Mismatches 76	Indels 5	Gaps 4

QY	64	GATGTTACTGTGACGGCCCCACAGTCGTATCTCGCCGAAAGGCGCGTCCCGTGT	123
Db	1	GATGCGTATTTGTGATGGCCCACTTGGCGTATTTACCGAAAGGCGCGTCCCGTGT	60
QY	124	AAAGTAATTCATTCGAAAGACGTCCGGGCGCGCGCGCTACGTGGCGATGAATATCGCT	183
Db	61	AAGGTAAATACCGTTGAGGAACGCCCGGCGCGCGCGGAACTGCGCATATACTTTCG	120
QY	184	TCTCTCGGTGCTAATGACACGCTGTGCGGGTGTGACGGGCAATTGACGATGACGCGCGCG	243
Db	121	TGCTGTGGAGCGAACCGCCGTCTGTGCGGCTGTACGGG-TTATGTATGACCCCGCGGCGC	179
QY	244	CTGAGTAAATCTCTGGCGCGACGTAAAGTGAAG-CGACTTCGTTCTGTAAACGACGCA	302
Db	180	CTGAGCAAAACGCTGGCGGAGGTCAATGTGAAGTGCAGACTTCGTTCTGTGCGGACGCA	239
QY	303	TCCGACATTAACCAATTACGGGTACTTTC-CGCGAACCAACAGCTGATCCGCTGTGATT	361
Db	240	TCCGACGATTAACCAATCTCGAGATCACTATCTAACGTAATCAGACGCTACTTGGTTG--AT	297
QY	362	TTGAAGAAAGGTTTGGAAAGTG	382
Db	298	TTGAAGAAAGGCTTTGAGGATG	318

```

/ RESULT 14
/ US-09-134-0368-185
/ Sequence 185, Application US/091940368
/ Patent No. 6548246
/ GENERAL INFORMATION:
/ APPLICANT: Mahan, Michael J.
/ Conner, Christopher P.
/ Hiehoff, Douglas W.
/ TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION
/ OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST
/ INFECTION
/ NUMBER OF SEQUENCES: 255
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morrison & Foerster
/ STREET: 755 Page Mill Road
/ CITY: Mountain View
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/

```



APPLICATION NUMBER: US/09/194, 036B  
FILING DATE: 17-NO. 6548246-1998  
CLASSIFICATION: <UNKNOWN>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/08208  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: US 08/651,155  
FILING DATE: 1996-05-17  
ATTORNEY/AGENT INFORMATION:  
NAME: Shantanu Basu  
REGISTRATION NUMBER: 43,318  
REFERENCE/DOCKET NUMBER: 22002060601  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 813-5995  
TELEFAX: (650) 494-0792  
INFORMATION FOR SEQ ID NO: 185:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: YES  
ORIGINAL SOURCE:  
ORGANISM: DNA (other)  
SEQUENCE DESCRIPTION: SEQ ID NO: 185:  
US-09-194-036B-185

Query Match 10.4%; Score 149.4; DB 3; Length 347;  
Best Local Similarity 74.8%; Pred. No. 1.7e-33;  
Matches 240; Conservative 0; Mismatches 76; Indels 5; Gaps 4;

QY 64 GATGTTACTGCTGACGCGCCCAACAGTCTGCGCGGAGCGCGCGTTCGTT 123  
DB 1 GATGCTATTGCTGATGCGCCCACTGCGCTATTTACCGGAAAGCGCGTCCGTT 60

QY 124 AATGTAATACATCGAAGACGTCCGGGCGCGCGCTACGTCGCGATGATGCT 183  
DB 61 AAGGTAATACGCTGTAAGAACGCGCGCGCGCGGAGAACGTCGCGATGATGCT 120

QY 184 TCTCTCGCTGCTAAATGACGCGCTGCTGCTGTAAGCGGCAATGACGACGCG 243  
DB 121 TGCCTGGAAGCGAACGCGCTGCTGCTGCTGCGCTGAAGCGG-TTATGATGAGCGCGCGCGC 179

QY 244 CTGAGTAATCTGCGCGCGACGTCAGCTCAATG-CGACTTCTGTTCTGACGAGCA 302  
DB 180 CTGAGCAAAAGCTGCGCGAGCTCAATGTAAGTCCGACTTCTGTCGCGAGCA 239

QY 303 TCGACCATTCACAAATTACGGGTACTTC-CGCGAACCAAGCTGATCCGCTGAT 361  
DB 240 TCCGACATTAACAACTGCGAGTACTATGTAATCAAGCAAGCTCATTCGTTG--AT 297

QY 362 TTGAAGAAGTTTCGAAGTG 382  
DB 298 TTGAAGAAGCTTTGAGGATG 318

RESULT 15  
US-09-252-991A-16179/c  
Sequence 16179, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 16179  
LENGTH: 789  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16179

Query Match 10.1%; Score 144.8; DB 3; Length 789;  
Best Local Similarity 58.3%; Pred. No. 5.3e-32;  
Matches 273; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

QY 1 ATGAAGTAACTGTCGACGAGTTTGAACGTGACGAGATGATGCTGATGATGATG 60  
DB 465 ATGAAGTTCATGACGCGCGTTTCGACCAAGCCCGGTGTTGATGCTGATGATGATG 406

QY 61 CTGATGTTACTGCTGACGCGCCCAACAGTCTGATCTGCGCGAAGCGCGTCCCTG 120  
DB 405 CTCGACCGCTATTGCGATGCGGAGCTTCGCGCAATTCGCGGAGCCCGGTGCGGTG 346

QY 121 GTTAAAGTAACTGTCGACGAGTTTGAACGTGACGAGATGATGCTGATGATGATG 180  
DB 345 GTCCGCTGCAACGACGAGACCGCCCGCGCGCGCGCGCAACGTCGCTGATGATG 286

QY 181 GCTTCTCTGCTGCTGATGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 285 GCGCGCTGCGCGCGGAGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 226

QY 241 GCGCTGATTAATCTGCGCGCGACGCTCAATGCAATGCAATGCTTCTTGTACGACG 300  
DB 225 AGCTGCGCAACAGCTCAAGCGCGCTGAGTGAACGCGCTTCCAGCGCATGATGAC 166

QY 301 CATCCGACATTAACAAATTACGGGTACTTCCGCAACCAAGCATCGCTGATGATG 360  
DB 165 CAGCGACATCTGCTCAAGCTGCGGTCATGATGCTGCTGCTGCTGCTGCTGCTGCTG 106

QY 361 TTGAAGAAGTTTCGAAGGTGTTGATCCGACCGCTGACGAGCGGATTAATGAGCG 420  
DB 105 TTGAGGAAACGTTT---CGCACGACGCGCGCGCGCTGCGATGAGTGTGCTG 49

QY 421 CTGAGTTGATGCGCGCTGCTGCTTCTGATCTGACGCGCAAGGTGCG 468  
DB 48 CTGCGCAAGGTCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1

Search completed: March 18, 2006, 19:21:03  
Job time: 469 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2006, 06:18:38 ; Search time 24 Seconds  
(without alignments)  
568,880 Million cell updates/sec

Title: US-10-520-820-13

Perfect score: 2393  
Sequence: 1 MKTLPFERBAGVWVGDVW.....PDCSTTNIIKKIQDKG 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 26622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications AA New:  
1: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2393	100.0	477	6	US-10-520-820-13
2	744.5	31.1	320	6	US-10-467-657-3254
3	480.5	20.1	345	7	US-11-096-666-11061
4	341.5	14.3	169	6	US-10-467-657-548
5	324.5	13.6	169	7	US-11-096-666-10264
6	229.5	9.6	323	7	US-11-132-864-7
7	183.5	7.7	361	7	US-11-096-568A-6281
8	183.5	7.7	369	7	US-11-096-568A-6281
9	181.5	7.6	378	7	US-11-096-568A-31212
10	181.5	7.6	379	7	US-11-096-568A-31211
11	181.5	7.6	705	7	US-11-096-568A-31090
12	181.5	7.6	741	7	US-11-096-568A-31089
13	181.5	7.6	750	7	US-11-096-568A-31088
14	166	6.9	261	7	US-11-096-568A-31213
15	164.5	6.9	212	7	US-11-096-568A-6282
16	154.5	6.5	302	6	US-10-793-626-3062
17	153	6.4	307	7	US-11-087-099-4308
18	152.5	6.4	323	7	US-11-096-568A-10265
19	152.5	6.4	359	7	US-11-096-568A-10265
20	144.5	5.8	323	7	US-11-087-099-1760
21	139.5	5.8	323	7	US-11-087-099-1760
22	132	5.5	316	7	US-11-087-099-9229
23	130.5	5.5	488	7	US-11-087-099-303
24	128	5.3	7968	7	US-11-143-980-49
25	125.5	5.2	544	6	US-10-467-657-1020

26	125.5	5.2	545	7	US-11-201-916-19	Sequence 19, Appl
27	123.5	5.2	335	7	US-11-087-099-1705	Sequence 1705, Ap
28	123.5	5.2	333	7	US-11-096-568A-31752	Sequence 31752, A
29	120.5	5.0	401	7	US-11-096-568A-34248	Sequence 34248, A
30	118	4.9	314	7	US-11-096-568A-31753	Sequence 31753, A
31	116.5	4.9	372	7	US-11-096-568A-31152	Sequence 31152, A
32	116.5	4.9	404	7	US-11-096-568A-31151	Sequence 31151, A
33	116.5	4.8	371	7	US-11-096-568A-34249	Sequence 34249, A
34	112	4.7	1571	7	US-11-052-554A-2	Sequence 2, Appl1
35	110.5	4.6	1121	7	US-11-087-099-8532	Sequence 8532, Ap
36	109	4.6	524	7	US-11-082-389-10	Sequence 10, Appl
37	108.5	4.5	335	7	US-11-087-099-4364	Sequence 4364, Ap
38	108	4.5	633	6	US-10-467-657-6628	Sequence 6628, Ap
39	107.5	4.5	325	7	US-11-087-099-7465	Sequence 7465, Ap
40	107	4.5	319	6	US-10-793-626-786	Sequence 786, Ap
41	107	4.5	352	6	US-10-840-688-22	Sequence 22, Appl
42	107	4.5	3507	7	US-11-075-185-7	Sequence 7, Appl1
43	106.5	4.5	537	6	US-10-467-657-4598	Sequence 4598, Ap
44	106.5	4.5	1430	7	US-11-175-689-9	Sequence 9, Appl1
45	106	4.4	319	7	US-11-087-099-880	Sequence 880, Appl

## ALIGNMENTS

RESULT 1  
US-10-520-820-13  
Sequence 13, Application US/10520820  
Publication No. US2006003393A1  
GENERAL INFORMATION:  
APPLICANT: MUTABILIS S.A.  
TITLE OR INVENTION: Pathogenicity determinants which can be used as targets for devel  
TITLE OR INVENTION: means for preventing and controlling bacterial infections and/or  
TITLE OR INVENTION: dissemination  
FILE REFERENCE: 1621  
CURRENT APPLICATION NUMBER: US/10/520, 820  
CURRENT FILING DATE: 2005-01-07  
PRIOR APPLICATION NUMBER: PCT/EP2003/008209  
PRIOR FILING DATE: 2003-07-09  
PRIOR APPLICATION NUMBER: FR 0208636  
PRIOR FILING DATE: 2002-07-09  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13  
LENGTH: 477  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-10-520-820-13

Query Match	100.0%	Score 2393;	DB 6;	Length 477;
Best Local Similarity	100.0%	Pred. No. 2.5e-164;		
Matches 477;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	MKTLPFERBAGVWVGDVWLDRTYWGPTSRISPEAPVPPVKNVTEERPGANVAMI	60		
1	MKTLPFERBAGVWVGDVWLDRTYWGPTSRISPEAPVPPVKNVTEERPGANVAMI	60		
61	ASIGANRLVGLTGIDDAARLSKSLADVANKCDPVSPTPTTKRLVSRNQQLRLD	120		
61	ASIGANRLVGLTGIDDAARLSKSLADVANKCDPVSPTPTTKRLVSRNQQLRLD	120		
121	FERGFEVDVDPPIHERINOMLSISIGLVSDYKAGLASVQOMIQLARKGVPLIDPKG	180		
121	FERGFEVDVDPPIHERINOMLSISIGLVSDYKAGLASVQOMIQLARKGVPLIDPKG	180		
181	TDPERYCATLTPNLSFEFAVVGKCTEERIEYERGWKLADYELSLVTRSEOGMSLL	240		
181	TDPERYCATLTPNLSFEFAVVGKCTEERIEYERGWKLADYELSLVTRSEOGMSLL	240		
241	QPKAPLHMPFOAQEVYDVYTGADTVYIGVLAATLAAGNSLEBAKCFPANAAGVVGKLG	300		
241	QPKAPLHMPFOAQEVYDVYTGADTVYIGVLAATLAAGNSLEBAKCFPANAAGVVGKLG	300		

Qy	301	STYSPIELENAVGRADTCGCVWTEBELTLAAARKGKGVMMNGYFDLLHGHVSYL	366
Db	301	STVSPITLENAVGRADTCGCVWTEBELTLAAARKGKGVMMNGYFDLLHGHVSYL	366
Qy	361	ANARKGDRIVAVNSDASTKRLKGSRPVNLBOPLVTLGALBAVDVWSFEEDTPQRL	420
Db	361	ANARKGDRIVAVNSDASTKRLKGSRPVNLBOPLVTLGALBAVDVWSFEEDTPQRL	420
Qy	421	IAGILPDLLKKGSDYKPEETAGSKBVMWANGSEVLVLANFEDGCSSTNIITKKIQDQKKG	477
Db	421	IAGILPDLLKKGSDYKPEETAGSKBVMWANGSEVLVLANFEDGCSSTNIITKKIQDQKKG	477

```

RESULT 2
US-10-467-657-3254
Sequence 3254, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMan99, version 1.04
SEQ ID NO 3254
LENGTH: 320
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3254

```

Query Match	Score	DB	Length
31.1%	744.5	6	320
Best Local Similarity	51.0%	Pred. No. 3.3e-46	
Matches 157; Conservative	48;	Mismatches 102;	Indels 1;
			Gaps 1

Oy		8	FEBAQVWVGDWMLDBYRWGPTSRISPEEA.PVVVVVNITTEBPGGAANVANNTISLGANA	67
D*		14	PAQAATLVAGDVMILDRMTFEDVNSRISEAPAVPAKIGRIDQAGAANVAANNISLGGRA	73
Oy		68	RIVELTGIDDAABALSKSLADVNVKKDPFVSVPHTPTTKRLVLSNQOLIRLDPEEGEG	127
Db		74	GILSVTDDDBAAAALLDALMWODGVASLYLRMDRKQIAITTVGLRVARNQOOLIRLDFEEHPNR	133
Oy		128	VDPQPLHERLNQALSSIGLVSDVKAGLASVOQMOLARAGAPVYLIDPKGNDFPFRYR	187
Db		134	EVLQIKRRKRRELLPEYDAILTFSDYGKGGLSHSIDMIPAKVGKTVALIDPKSGDYETKYV	193
Oy		188	GATILLTNLISEFAVNGKCTEIBEIVERGMKLIALDEYSALLVTRECOGMSLIOPGRAPL	247
Db		194	GATLITNCAMELBKVWGSMNEGDLITEKQNRIHRHIDLTAVALLTIREBGMTLPFBEGH-PI	252
Oy		248	HMPFOAEVVDYMGAGGYTVIGVLAATLAAAGNLSIEACFPANAAAAVWVGKLGSTSVSPIS	307
Db		253	XQPRABBYVDVSGAGDTYIAGCGLOAAGCTMPBAMTYANTAAAGVVAKAGTAVCSFAE	312
Oy		308	LENAVRGR	315
Db		313	LVEALDGO	320

RESULT 3  
US-11-098-686-11061  
; Sequence 11061, Application US/11098686  
; Publication NO. US20060024696A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES

```

; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
;
; FILE REFERENCE: 09531-128001
;
; CURRENT APPLICATION NUMBER: US/11/098,686
;
; CURRENT FILING DATE: 2005-04-04
;
; PRIOR APPLICATION NUMBER: PCT/US03/31318
;
; PRIOR FILING DATE: 2003-10-01
;
; PRIOR APPLICATION NUMBER: US 60/416,395
;
; PRIOR FILING DATE: 2002-10-04
;
; NUMBER OF SEQ ID NOS: 11433
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 11061
;
; LENGTH: 345
;
; TYPE: PRF
;
; ORGANISM: Lawsonia intracellularis
;
; US-11-098-686-11061

```

Query Match	20.1%	Score 480.5;	DB 7;	Length 345;
Best Local Similarity	37.0%	Pred. No. 3.3e-27;		
Matches 117; Conservative	51;	Mismatches 127;	Indels 21;	Gaps 6;

```

QY 13 MMVGGWMLDHYRKYGPTSRISPEAPVPVKNUTTEERGGANVAMNITSLANRLVBL 72
Db 23 VLVGGIDMLDNYLIVSDRISPEAPVPLVKNENQSGAGNVARSLALGKXITIIGA 62
QY 73 TGIDDAAPALSKSLADVAVKCDPVSVP7PHTTKLVL7SRNQQILRDPEE--GFEQVDP 130
Db 83 VGOQDSGSKIIDLLSTRGILSIIITFANRQTVTKTRVWAHRQOMRLDHEESTVNSKEL 142
QY 131 QPLHERIQALSSIGALVLDYAKALASVQ-----QMQLQARKAGVPLIDPKGDPEER 185
Db 143 AMVLSNFEKYSQHHIIILSDYNGK-LVSKERMLGFOQILILAKNNNAKYLIDRPGNIY 201
QY 186 Y---RGATLLLPNTSEFAVNVGKC-----KTEERIYERGMKLIADYELISALLVTRSEOG 236
Db 202 YALCNIPALIPENTKE---TGECMGGMATSSQPELLAAGHTIMKLSTKHLITTYGDSG 257
QY 237 MSLLQPKAPLHPMPAQOEVDVDTGAGDTVIGVLATLAAGNSLEBACEFFANAAAGVVVG 296
Db 258 MALFLSPKKIHHIPIVGRDVFDTVGAAGDTVLATPELIALSAGLDPLISALIVAAAGVVVS 317
QY 297 KLGTSTVSPFLENNAV 312
Db 318 QVGTATVSPDELEBAI 333

```

```

RESULT 4
US-10-467-657-548
; Sequence 548, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqwIn99, version 1.04
; SEQ ID NO 548
; LENGTH: 169
; TYPE: FRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-548

```

Query Match	14.3%;	Score 341.5;	DB 6;	Length 169;
Best Local Similarity	51.8%;	Pred. No. 1.1e-17;		
Matches 71;	Conservative 21;	Mismatches 44;	Indels 1;	Gaps 1;

Qy 342 VMTNGVFDILHAGHVSILANARKLGDRLIVAVNSDASTKRL-KGDSRPVPLEQRMTVL 400  
 Db 32 LVFTNGCPDILHKGHTVTLQARSGALVLAITLTDASVRLGSGDRPNVPLENRALVA 91  
 Qy 401 GALEAVDVMVFEEDTQRLIAGILPDLLVKGDKPEEISAGSEKXWANGSEVLVLPED 460  
 Db 92 AALESVDLVMPEDDTFRLALIEAVKPEVLVKGDMVVVDKIVGAETILARGGVPSIFLH 151  
 Qy 461 GCSTTNIKKIQODKKG 477  
 Db 152 QTSSTTKIARIRAEKG 168

RESULT 5  
 US-11-098-686-10264  
 / Sequence 10264, Application US/11098686  
 / Publication No. US20060024696A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Kapur, Vivek and Gebhart, Conle J.  
 / TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
 / TITLE OF INVENTION: FROM LAMSONIA INTRACELLULARIS AND METHODS OF USING  
 / FILE REFERENCE: 09531-128001  
 / CURRENT APPLICATION NUMBER: US/11/098, 686  
 / PRIOR FILING DATE: 2005-04-04  
 / PRIOR APPLICATION NUMBER: PCT/US03/31318  
 / PRIOR FILING DATE: 2003-10-01  
 / PRIOR APPLICATION NUMBER: US 60/416,395  
 / PRIOR FILING DATE: 2002-10-04  
 / NUMBER OF SEQ ID NOS: 11433  
 / SOFTWARE: FASTSEQ for Windows Version 4.0  
 / SEQ ID NO 10264  
 / LENGTH: 169  
 / TYPE: PRT  
 / ORGANISM: Lawsonia intracellularis  
 / US-11-098-686-10264

Query Match 13.6%; Score 324.5; DB 7; Length 169;  
 Best Local Similarity 46.1%; Pred. No. 1.9e-16;  
 Matches 65; Conservative 27; Mismatches 48; Indels 1; Gaps 1;  
 Qy 336 RKGGEKVTMTNGVFDILHAGHVSILANARKLGDRLIVAVNSDASTKRL-KGDSRPVPLE 394  
 Db 25 KSNKGKIIIFNCGVDILHGHLDLTAKTYGDLIVGLMTDVSVKRLGDPDRFPPFH 84  
 Qy 395 QRMTVLGALAEVAVNSFEEDTQRLIAGILPDLLVKGDKPEEISAGSEKXWANGSEVL 454  
 Db 85 IRAFVLALHLELDIFIEFEEDTQRLIAGILPDLLVKGDKPEEISAGSEKXWANGSEVL 144  
 Qy 455 VLNEDGCGSTTNIKKIQODK 475  
 Db 145 SLPELEGHSTSGLVOKIRNKK 165

RESULT 6  
 US-11-132-864-7  
 / Sequence 7, Application US/11132864  
 / Publication No. US20050289670A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Jintul Shi  
 / APPLICANT: David Erci  
 / APPLICANT: Lisa Hagen  
 / APPLICANT: Hongyu Wang  
 / TITLE OF INVENTION: Plant Myo-Inositol Kinase  
 / TITLE OF INVENTION: Polynucleotides and Methods of Use  
 / FILE REFERENCE: 035718/291638  
 / CURRENT APPLICATION NUMBER: US/11/132, 864  
 / PRIOR FILING DATE: 2005-05-19  
 / PRIOR APPLICATION NUMBER: 60/573,000  
 / PRIOR FILING DATE: 2004-05-20  
 / NUMBER OF SEQ ID NOS: 48  
 / SOFTWARE: FASTSEQ for Windows Version 4.0  
 / SEQ ID NO 7  
 / LENGTH: 323

TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Pfam consensus for pfkb family  
 US-11-132-864-7

Query Match 9.6%; Score 229.5; DB 7; Length 323;  
 Best Local Similarity 29.2%; Pred. No. 3.1e-09;  
 Matches 102; Conservative 35; Mismatches 127; Indels 85; Gaps 14;  
 Qy 13 VMTNGVMDLRYVYGTSPRISPEAPVVKVNTIERPGGA-ANVAMNIALSIG-ANA 67  
 Db 2 VVVIGANMDLI-PTVGLPEGEIN-ERYVSFEKPGAGANVVALARIGNPSYGV 57  
 Qy 68 RLVLGTGIDDAARALSKSLADVVKCDPVSVPT-HPITTKLRVUS 111  
 Db 58 APFGKVDDEFGFELLLKKEGVVDVYVNGEGRGTGLALVLPDGGERRITVPRGAN 117  
 Qy 112 RNOQLRLDPEBEGVD-POPLHERINQALSSIGALVLSYAKGALAS 159  
 Db 118 ADTLTSLD-EDLLEADILHSGISLVLLPELPETLEALAE-160  
 Qy 160 VQOMIOLARKAGVPLIDPKTD-PEYRGATLTPNLSPEPAVVGCK 207  
 Db 161 -AAKAGKISFDPLNDPLMSDEALVELLEPLADILKPEEELTLGK 213  
 Qy 208 TEEIYERGMKLIADYELSL-LVTRSEQMSLQPGKAPLHPTQAO-EYDVTGAGDT 265  
 Db 214 EVERBALAHLKILAKAVTKLVVVTGADGALLVTG-GEVHPAPVKVVDVTGAGDA 272  
 Qy 266 -VIGVLAATLAA-GNSLEACFPANAAAGVVGKIGTSVSP 305  
 Db 273 FVAGFLAGLITDSTQLDGKDLERLRFANAAALVQKKAISLP 321

RESULT 7  
 US-11-096-568A-6281  
 / Sequence 6281, Application US/11096568A  
 / Publication No. US20060048240A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Alexandrov, Nikolai et al.  
 / TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 / TITLE OF INVENTION: Thedy  
 / FILE REFERENCE: 2750-1592PUS2  
 / CURRENT APPLICATION NUMBER: US/11/096, 568A  
 / PRIOR FILING DATE: 2005-04-01  
 / NUMBER OF SEQ ID NOS: 34471  
 / SEQ ID NO 6281  
 / LENGTH: 361  
 / TYPE: PRT  
 / ORGANISM: Glycine max  
 / FEATURE:  
 / NAME/KEY: misc feature  
 / LOCATION: (1)-(361)  
 / OTHER INFORMATION: Ceres Seq. ID no. 14314595  
 US-11-096-568A-6281

Query Match 7.7%; Score 183.5; DB 7; Length 361;  
 Best Local Similarity 27.0%; Pred. No. 7.4e-06;  
 Matches 85; Conservative 41; Mismatches 146; Indels 43; Gaps 12;  
 Qy 13 VMTNGVMDLRYVYGTSPRISPEAPVVKVNTIERP-GGANVAMNIALSIGAN 66  
 Db 56 VVVVGSANMD-LYEVNDRIPGEG-ETLARSGQTLAGGAGVQATCSALAP 106  
 Qy 67 ARVLGTGIDDAARALSKSLADVNVKCD-FVSVPTHTITKRLVLSRNOQLILDE 122  
 Db 107 TYVVGVDYDAYRVLTYAGRGCVRLDNLAVASATGHAVVMLQSNQNSIVY-161  
 Qy 123 EGFEVD-POPLHERINQALSSIGALVLSYAKGALASVQOMIOLARKAGVPLID 178  
 Db 162 -IGANTLSCWPSLPROHLIDVAQAGIVLQREIRIDAVVA-QVAQAANKAGVVPVLA 217

QY 179 KGTDF----ERYRGATLTPNLSEFEAVG-KCTEEIEVERGKLIADYELSA--LLVT 231  
 Db 218 GGNMGPRLPQINIVNDLSPNETLRLTGMPTSEFEEIOALKC---HEMGAKQVLVK 274  
 QY 232 RSEOGMSLLOPGKAPLHMP- QAOEYVDVTGAGDTVIGVLAATLAAGNSLEBACFPANAA 290  
 Db 275 LGHKGSLPVEGEKTIQQPALIAKTVDTTGAGDTFPAAFVAVLVEGSKSKECIRFAPAAA 334  
 QY 291 AGVVVGKLGSTVSP 305  
 Db 335 ACLCVQVKASPSMP 349

## RESULT 8

US-11-096-568A-6280  
 ; Sequence 6280, Application US/11096568A  
 ; Publication No. US20060048240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexandrov, Nickolai et al.  
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 ; TITLE OF INVENTION: Thierby  
 ; FILE REFERENCE: 2750-1592PUS2  
 ; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; CURRENT FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471  
 ; SEQ ID NO 6280  
 ; LENGTH: 369  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)..(369)  
 ; OTHER INFORMATION: Ceres Seq. ID no. 14314594  
 US-11-096-568A-6280

Query Match 7.7%; Score 183.5; DB 7; Length 369;  
 Best Local Similarity 27.0%; Pred. No. 7.7e-06;  
 Matches 85; Conservative 41; Mismatches 146; Indels 43; Gaps 12;

QY 13 VNVVGDVMDLRVYGPTRISPEAPVVPVVKNTIEERP-----GGANVAMNITASIGAN 66  
 Db 64 VVVGSANAD--IYEVDRLEPEG-----ETLAASGOTIAGKGANQATCSAKIAYP 114  
 QY 67 ARLVGLTGIDDAARALSKSLADVAVKCD---FVSVPPTHPTITLRLVLSRNOQLIRLDFE 122  
 Db 115 TYTFGVQVGDADAYGSLVTRAGLRGGVRLDNLAVASATGAHVWVLQSGNCSIYV----- 169  
 QY 123 EGFEQVD---PQPLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDP 178  
 Db 170 --IGGANLSCWPSLPRQHLDLVAQAGIVLLQREIPDAVNA--QVAQAAKNAAGVPLVLA 225  
 QY 179 KGTDF----ERYRGATLTPNLSEFEAVG-KCTEEIEVERGKLIADYELSA--LLVT 231  
 Db 226 GGNMGPRLPQINIVNDLSPNETLRLTGMPTSEFEEIOALKC---HEMGAKQVLVK 282  
 QY 232 RSEOGMSLLOPGKAPLHMP- QAOEYVDVTGAGDTVIGVLAATLAAGNSLEBACFPANAA 290  
 Db 283 LGHKGSLPVEGEKTIQQPALIAKTVDTTGAGDTFPAAFVAVLVEGSKSKECIRFAPAAA 342  
 QY 291 AGVVVGKLGSTVSP 305  
 Db 343 ACLCVQVKASPSMP 357

RESULT 9  
 US-11-096-568A-31212  
 ; Sequence 31212, Application US/11096568A  
 ; Publication No. US20060048240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexandrov, Nickolai et al.  
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 ; TITLE OF INVENTION: Thierby  
 ; FILE REFERENCE: 2750-1592PUS2

; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; CURRENT FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471  
 ; SEQ ID NO 31212  
 ; LENGTH: 378  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)..(378)  
 ; OTHER INFORMATION: Ceres Seq. ID no. 13579029  
 US-11-096-568A-31212

Query Match 7.6%; Score 181.5; DB 7; Length 378;  
 Best Local Similarity 25.6%; Pred. No. 1.1e-05;  
 Matches 81; Conservative 45; Mismatches 142; Indels 49; Gaps 11;

QY 28 PYSRISPEAPVVPVVKNT-----IEERP-----GGANVAMNITASIGAN 63  
 Db 59 PKSAVDANHP-PLVVVGSANADIYVEIERLPKEGETISAKTGOTLAGKGANQACGAKL 117  
 QY 64 GANARLVGLTGIDDAARALSKSLAD--VAVKCDV-SVPTHPT---ITLRLVLSRNOQL 117  
 Db 118 MYPTFVGRLGSDAHGKLIAGALGDDCGVHLDYRVSNNBPFGAAVWVLQSDGNSI 177  
 QY 118 RLDPEEGFEQVD---PQPLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGV 173  
 Db 178 -----VGANWKMPELMSDDLEIVNAGIVLLQREIPDSIN--QVAARVKGAGVP 228  
 QY 174 VLIDPKGTDF---ERYRGATLTPNLSEFEAVGKCTEEIEVERGKLIADYELSA 229  
 Db 229 VILDVGMDPTLPENLIDSIDLSPNETLRLTGMPTSEFEEIOALKC---HEMGAKQVLVK 282  
 QY 230 VRSBEGMSLLOPGKAPLHMP- QAOEYVDVTGAGDTVIGVLAATLAAGNSLEBACFPAN 288  
 Db 289 VKLSGKSLPFIQGEKPIQOSIIPAAQVVDVTGAGDTFPAAFVAVLVEGSKSKECIRFAP 348  
 QY 289 AAGVVVGKLGSTVSP 305  
 Db 349 AASLVCVQVKALPSMP 365

RESULT 10  
 US-11-096-568A-31211  
 ; Sequence 31211, Application US/11096568A  
 ; Publication No. US20060048240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexandrov, Nickolai et al.  
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 ; TITLE OF INVENTION: Thierby  
 ; FILE REFERENCE: 2750-1592PUS2  
 ; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; CURRENT FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471  
 ; SEQ ID NO 31211  
 ; LENGTH: 379  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)..(379)  
 ; OTHER INFORMATION: Ceres Seq. ID no. 13579028  
 US-11-096-568A-31211

Query Match 7.6%; Score 181.5; DB 7; Length 379;  
 Best Local Similarity 25.6%; Pred. No. 1.1e-05;  
 Matches 81; Conservative 45; Mismatches 142; Indels 49; Gaps 11;

QY 28 PYSRISPEAPVVPVVKNT-----IEERP-----GGANVAMNITASIGAN 63  
 Db 60 PKSAVDANHP-PLVVVGSANADIYVEIERLPKEGETISAKTGOTLAGKGANQACGAKL 118  
 QY 64 GANARLVGLTGIDDAARALSKSLAD--VAVKCDV-SVPTHPT---ITLRLVLSRNOQL 117



Db 119 MYPTFYGRIGEDAHGGLIALBALDCCGCHLDVVRVNNPETHAIVMLQSDGQNSIII 178  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 118 RLDPEBEGVD---POLHERINQALSSIGALVSDYAKGALASVOQMQLARKKAVP 173  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 179 -----VGGANKMKAPEIMSDDELEIVRNAGIVLQREIPDSINI--QVAKAVKKGAVP 229  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 174 VLIDPKGTDF-----ERRRGATLLTPNLSEPEAVVVGKCTEERIVERGMKLIADVEISALL 229  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 230 VILVGGMDTPPIPNELIDSLIDILSPNETELSRLTGMTETFEQISQVAKCHKLGKQVL 289  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 230 VTRSEQMSLQPGKAPLHNP-TOAQEVYDVYTGAGDTVIGVLAATLAAGNSLEBACFPAN 288  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 290 VKGSGKSALFIQGEKPIQOSIIPAAQVVDVTGAGDTFTAFAFNAVMEGKSHEBCLAFPA 349  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 289 AAGVYVVGKLGSTVSP 305  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 350 AAASLCTQVKGALPSPMP 366  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

## RESULT 11

US-11-096-568A-31090  
; Sequence 31090, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 31090  
; LENGTH: 705  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(705)  
; OTHER INFORMATION: Ceres Seq. ID no. 4987030  
US-11-096-568A-31090

Query Match 7.6%; Score 181.5; DB 7; Length 705;

Best Local Similarity 25.6%; Pred. No. 2.7e-05;

Matches 81; Conservative 45; Mismatches 142; Indels 49; Gaps 11;

Qy 28 PFSRISPEAPVPVVKNT-----IEERP-----GGAANYAMNIASL 63  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 386 PKSAVDNHP-PLVVGSANADIVYEIERLPKEGETISAKTGQTLAGKGANQAAAGAKL 444  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 64 GANARLVGLTGIDDAARALSKSLAD--VWYKCDPV-SVPTHTPT---ITKLRVLSRNOQLI 117  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 445 MYPTFYGRIGEDAHGGLIALBALDCCGCHLDVVRVNNPETHAIVMLQSDGQNSIII 504  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 118 RLDPEBEGVD---POLHERINQALSSIGALVSDYAKGALASVOQMQLARKKAVP 173  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 505 -----VGGANKMKAPEIMSDDELEIVRNAGIVLQREIPDSINI--QVAKAVKKGAVP 555  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 174 VLIDPKGTDF-----ERRRGATLLTPNLSEPEAVVVGKCTEERIVERGMKLIADVEISALL 229  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 230 VILVGGMDTPPIPNELIDSLIDILSPNETELSRLTGMTETFEQISQVAKCHKLGKQVL 289  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 230 VTRSEQMSLQPGKAPLHNP-TOAQEVYDVYTGAGDTVIGVLAATLAAGNSLEBACFPAN 288  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 616 VKGSGKSALFIQGEKPIQOSIIPAAQVVDVTGAGDTFTAFAFNAVMEGKSHEBCLAFPA 675  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 289 AAGVYVVGKLGSTVSP 305  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 676 AAASLCTQVKGALPSPMP 692  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

## RESULT 12

US-11-096-568A-31089

; Sequence 31089, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 31089  
; LENGTH: 741  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(741)  
; OTHER INFORMATION: Ceres Seq. ID no. 4987029  
US-11-096-568A-31089

Query Match 7.6%; Score 181.5; DB 7; Length 741;

Best Local Similarity 25.6%; Pred. No. 2.9e-05;

Matches 81; Conservative 45; Mismatches 142; Indels 49; Gaps 11;

Qy 28 PFSRISPEAPVPVVKNT-----IEERP-----GGAANYAMNIASL 63  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 422 PKSAVDNHP-PLVVGSANADIVYEIERLPKEGETISAKTGQTLAGKGANQAAAGAKL 460  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 64 GANARLVGLTGIDDAARALSKSLAD--VWYKCDPV-SVPTHTPT---ITKLRVLSRNOQLI 117  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 481 MYPTFYGRIGEDAHGGLIALBALDCCGCHLDVVRVNNPETHAIVMLQSDGQNSIII 540  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 118 RLDPEBEGVD---POLHERINQALSSIGALVSDYAKGALASVOQMQLARKKAVP 173  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 541 -----VGGANKMKAPEIMSDDELEIVRNAGIVLQREIPDSINI--QVAKAVKKGAVP 591  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 174 VLIDPKGTDF-----ERRRGATLLTPNLSEPEAVVVGKCTEERIVERGMKLIADVEISALL 229  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 592 VILVGGMDTPPIPNELIDSLIDILSPNETELSRLTGMTETFEQISQVAKCHKLGKQVL 651  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 230 VTRSEQMSLQPGKAPLHNP-TOAQEVYDVYTGAGDTVIGVLAATLAAGNSLEBACFPAN 288  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 652 VKGSGKSALFIQGEKPIQOSIIPAAQVVDVTGAGDTFTAFAFNAVMEGKSHEBCLAFPA 711  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Qy 289 AAGVYVVGKLGSTVSP 305  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
Db 712 AAASLCTQVKGALPSPMP 728  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

## RESULT 13

US-11-096-568A-31088  
; Sequence 31088, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 31088  
; LENGTH: 750  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(750)  
; OTHER INFORMATION: Ceres Seq. ID no. 4987028  
US-11-096-568A-31088

Query Match 7.6%; Score 181.5; DB 7; Length 750;

Best Local Similarity 25.6%; Pred. No. 3e-05;

Matches 81; Conservative 45; Mismatches 142; Indels 49; Gaps 11;

```

QY 28 PLSRISSEAPVYVYKVT-----IEERP-----GGANVAMNTASL 63
DB 431 PKAVNDHAP-PLVVGSSANADIYVEIRLPKSEETISAKTGQTLAGSKANQACAKL 489
QY 64 GANARVLGTGIDDAARALSKSLAD--VNYKCDPV-SVPTHPT---ITKLRVLSRNOQLI 117
DB 490 MYPLFYVGRIGEDAHGKLIABALDGCYHLDYRSVNNPEPTGHAVVLMQSDQNSIIL 549
QY 118 RLDEEPEGEVD---PQLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKGV 173
DB 550 -----VGGANMKAMPEISDDLEIVRNAGIYVLQREIPDSINT--QVAKAVKAGVP 600
QY 174 VLIDPKGTDF---ERRRGATLTPNLSEFEAVVKKCTEERIYERGMKLIADYELASL 229
DB 601 VILVGGMDPIPNELLDSDILSPNETELSRLTGMPTFEQISQAVAKCHKLGVKQVL 660
QY 230 VTRSEQGSLLQPGKAPLHMP-TOAQEYVDVTGAGDVTIGVLAATLAAGNSLEBACFPAN 288
DB 661 VKGSKGSALFIQGEKPIQOSIIPAAQVVDVTGAGDTFTAAFAVAVMGSKSHECLAFPA 720
QY 289 AAAGVYVVKLGSTVSP 305
DB 721 AAASLCVQVKALPSMP 737

```

## RESULT 14

```

US-11-096-568A-31213
; Sequence 31213, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 31213
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(261)
; OTHER INFORMATION: Ceres Seq. ID no. 13579030
US-11-096-568A-31213

```

Query Match 6.9%; Score 166; DB 7; Length 261;

Best Local Similarity 25.9%; Pred. No. 8.4e-05;

Matches 65; Conservative 42; Mismatches 120; Indels 24; Gaps 8;

```

QY 70 VGLTGIDDAARALSKSLAD--VNYKCDPV-SVPTHPT---ITKLRVLSRNOQLIRLDP 123
DB 7 VGRIGEDAHGKLIABALDGCYHLDYRSVNNPEPTGHAVVLMQSDQNSIIL----- 60
QY 124 GFEGVD---PQLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKGVPLIDPK 179
DB 61 -VGGANMKAMPEISDDLEIVRNAGIYVLQREIPDSINT--QVAKAVKAGVPIIDVG 117
QY 180 GTDP-----ERRRGATLTPNLSEFEAVVKKCTEERIYERGMKLIADYELASLITRS 235
DB 118 GMDPIPNELLDSDILSPNETELSRLTGMPTFEQISQAVAKCHKLGVKQVLVKGSK 177
QY 236 GMSLLQPGKAPLHMP-TOAQEYVDVTGAGDVTIGVLAATLAAGNSLEBACFPANAA 294
DB 178 GSALFIQGEKPIQOSIIPAAQVVDVTGAGDTFTAAFAVAVMGSKSHECLAFPA 237
QY 295 VVKLGSTVSP 305
DB 238 VQVKALPSMP 248

```

## RESULT 15

```

US-11-096-568A-6282
; Sequence 6282, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6282
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(212)
; OTHER INFORMATION: Ceres Seq. ID no. 14314596
US-11-096-568A-6282

```

Query Match 6.9%; Score 164.5; DB 7; Length 212;

Best Local Similarity 33.6%; Pred. No. 7.9e-05;

Matches 51; Conservative 23; Mismatches 67; Indels 11; Gaps 5;

```

QY 162 QMIQLARKGVPLIDPKGTDF---ERRRGATLTPNLSEFEAVVGG-KCTEERIYER 216
DB 52 QVAAAKKAGVPLVDAGGMDGPLPPQLINPVDILSPNETELARLTGMPTFEETIOAA 111
QY 217 MKLIADYELSA--LVTRSEQGSLLQPGKAPLHMP-TOAQEYVDVTGAGDVTIGVLA 273
DB 112 LKC---HEMGAKQVVLKLGKGSALFVEGEKTIQPAIILAKTVVDVTGAGDTFTAAFAVA 168
QY 274 LAAGNSLEBACFPANAAAGVYVVKLGSTVSP 305
DB 169 LVGSKSKCECLRFPAAMAACTCVQVKGASPSMP 200

```

Search completed: March 18, 2006, 06:21:47

Job time : 25 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2006, 06:17:53 ; Search time 168 Seconds

(Without alignments)  
1186.337 Million cell updates/sec

Title: US-10-520-820-13

Perfect score: 2393

Sequence: 1 MKYTLPEFERAGVWVGDVW.....FEDGCTNLIKIKIQDQDKG 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA Main:\*

- 1: /cgn2\_6/ptodate/1/pubpaa/US07\_PUBCOMB.pap:\*
- 2: /cgn2\_6/ptodate/1/pubpaa/US08\_PUBCOMB.pap:\*
- 3: /cgn2\_6/ptodate/1/pubpaa/US09\_PUBCOMB.pap:\*
- 4: /cgn2\_6/ptodate/1/pubpaa/US10A\_PUBCOMB.pap:\*
- 5: /cgn2\_6/ptodate/1/pubpaa/US10B\_PUBCOMB.pap:\*
- 6: /cgn2\_6/ptodate/1/pubpaa/US11\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2393	100.0	477	US-09-741-669-469	Sequence 469, App
2	2393	100.0	477	US-09-912-020-325	Sequence 325, App
3	2393	100.0	477	US-09-815-242-10301	Sequence 10301, A
4	2393	100.0	477	US-10-282-122A-42614	Sequence 42614, A
5	2393	100.0	477	US-10-771-241-325	Sequence 325, App
6	2294	95.9	476	US-10-282-122A-56202	Sequence 56202, A
7	2265	94.7	477	US-09-815-242-13759	Sequence 13759, A
8	2265	94.7	477	US-10-282-122A-75478	Sequence 75478, A
9	2227	93.1	477	US-10-282-122A-60277	Sequence 60277, A
10	2196	91.8	463	US-10-282-122A-73168	Sequence 73168, A
11	2076	86.8	476	US-10-282-122A-78165	Sequence 78165, A
12	2007	83.9	474	US-10-282-122A-68720	Sequence 68720, A
13	1712	71.5	476	US-10-282-122A-58533	Sequence 58533, A
14	1701	71.1	476	US-10-282-122A-67124	Sequence 67124, A
15	1357.5	56.7	474	US-10-282-122A-69312	Sequence 69312, A
16	1350.5	56.4	474	US-09-815-242-12079	Sequence 12079, A
17	1350.5	56.4	474	US-10-282-122A-66769	Sequence 66769, A
18	1338.5	55.9	469	US-10-282-122A-67849	Sequence 67849, A
19	863	36.1	461	US-10-282-122A-54522	Sequence 54522, A
20	825	34.5	334	US-10-282-122A-49479	Sequence 49479, A
21	809.5	33.8	461	US-10-282-122A-58820	Sequence 58820, A
22	774.5	32.4	447	US-10-335-977-5564	Sequence 5564, App
23	745.5	31.2	323	US-10-282-122A-65852	Sequence 65852, A
24	745.5	31.2	323	US-10-275-026A-52	Sequence 52, App1
25	744.5	31.1	320	US-10-282-122A-65224	Sequence 65224, A
26	734.5	30.7	424	US-10-335-977-5563	Sequence 5563, App
27	732.5	30.6	313	US-10-282-122A-50827	Sequence 50827, A

28	691.5	28.9	316	4	US-10-282-122A-47619	Sequence 47619, A
29	679.5	28.4	328	4	US-10-282-122A-50163	Sequence 50163, A
30	564.5	23.6	315	4	US-10-369-493-37	Sequence 37, App1
31	420	17.6	315	4	US-10-369-493-10214	Sequence 10214, A
32	362	15.1	161	4	US-10-282-122A-50471	Sequence 50471, A
33	350	14.6	161	4	US-10-282-122A-46155	Sequence 46155, A
34	344	14.4	178	4	US-10-282-122A-49730	Sequence 49730, A
35	341.5	14.3	168	4	US-10-282-122A-65242	Sequence 65242, A
36	340.5	14.2	168	4	US-10-282-122A-65689	Sequence 65689, A
37	332	13.9	162	4	US-10-282-122A-50947	Sequence 50947, A
38	275	11.5	136	4	US-10-282-122A-62085	Sequence 62085, A
39	254.5	10.6	293	4	US-10-369-493-16604	Sequence 16604, A
40	218	9.1	302	5	US-10-501-282-3560	Sequence 3560, App
41	218	9.1	322	5	US-10-501-282-3562	Sequence 3562, App
42	211	8.8	362	4	US-10-424-599-175504	Sequence 175504, A
43	201.5	8.4	319	4	US-10-369-493-17260	Sequence 17260, A
44	199.5	8.3	300	4	US-10-369-493-13944	Sequence 13944, A
45	197	8.2	308	3	US-09-815-242-11854	Sequence 11854, A

## ALIGNMENTS

RESULT 1	
US-09-741-669-469	
Sequence 469, Application US/09741669	
Patent No. US2002022718A1	
GENERAL INFORMATION:	
APPLICANT: Forayth, R. Allym	
APPLICANT: Ohlsen, Karl L.	
APPLICANT: Zykkind, Judith W.	
TITLE OR INVENTION: Gene identified as required for	
TITLE OR INVENTION: proliferation of B. coli	
FILE REFERENCES: ELITRA.009A	
CURRENT APPLICATION NUMBER: US/09/741,669	
CURRENT FILING DATE: 2000-12-19	
PRIOR APPLICATION NUMBER: US 60/173005	
PRIOR FILING DATE: 1999-12-23	
NUMBER OF SEQ ID NOS: 481	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 469	
LENGTH: 477	
TYPE: PRT	
ORGANISM: Escherichia coli	
US-09-741-669-469	
Query Match	100.0%; Score 2393; DB 3; Length 477;
Best Local Similarity	100.0%; Pred. No. 1.3e-198;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKYTLPEFERAGVWVGDVWLDKRYTGPTSRISPEAPVPPVKKNTTBERRGANVAMNI 60
DB	1 MKYTLPEFERAGVWVGDVWLDKRYTGPTSRISPEAPVPPVKKNTTBERRGANVAMNI 60
QY	61 ASIGANARIVGLGIDDAARALSKSLADVYKCDPVSPPHTITLRLVSRNOQLRLD 120
DB	61 ASIGANARIVGLGIDDAARALSKSLADVYKCDPVSPPHTITLRLVSRNOQLRLD 120
QY	121 FREGFGEVDPPPLERINQALSSIGALVLDVAKGALASVOQMIOLARKAGVPLIDPKG 180
DB	121 FREGFGEVDPPPLERINQALSSIGALVLDVAKGALASVOQMIOLARKAGVPLIDPKG 180
QY	181 TDERRRGATLTPNLSPEFAVVGKCTBERIYERGMKLIADVELSALVTRBEGQSLL 240
DB	181 TDERRRGATLTPNLSPEFAVVGKCTBERIYERGMKLIADVELSALVTRBEGQSLL 240
QY	241 QPGKAPLHMPYTOQOBYVDVGAADTVYGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
DB	241 QPGKAPLHMPYTOQOBYVDVGAADTVYGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
QY	301 STVSPLELNAVAGRADTGFVMTBEELKLAVAARRGKGVMTNGVFDILHAGVSYL 360
DB	301 STVSPLELNAVAGRADTGFVMTBEELKLAVAARRGKGVMTNGVFDILHAGVSYL 360

Qy 361 ANARKDRLIVAVNSDASTKRLKDSRPVNPPLBORMIVLGALEAVDPVWVSFEEDTPORL 420  
Db 361 ANARKDRLIVAVNSDASTKRLKDSRPVNPPLBORMIVLGALEAVDPVWVSFEEDTPORL 420  
Qy 421 IAGILPDLVLKGGDYKPEEIASKEVWANGGEVLVNFEDGCSSTNNIKKIQQDKG 477  
Db 421 IAGILPDLVLKGGDYKPEEIASKEVWANGGEVLVNFEDGCSSTNNIKKIQQDKG 477

RESULT 2  
US-09-912-020-325  
Sequence 325, Application US/09912020  
Patent No. US20020045592A1  
GENERAL INFORMATION:  
APPLICANT: Zykend, Judith  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Trawick, John  
APPLICANT: Forsyth, R. Allen  
APPLICANT: Froelich, Jamie M.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: ESCHERICHIA COLI  
FILE REFERENCE: ELITRA.001DV1  
CURRENT APPLICATION NUMBER: US/09/912,020  
PRIOR FILING DATE: 2001-07-23  
PRIOR APPLICATION NUMBER: 09/492,709  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/117,405  
PRIOR FILING DATE: 1999-01-27  
NUMBER OF SEQ ID NOS: 485  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 325  
LENGTH: 477  
TYPE: PRT  
ORGANISM: E. Coli  
US-09-912-020-325

Query Match 100.0%; Score 2393; DB 3; Length 477;  
Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MKVTLPEFERAGVWVGVDMLDRIYVGPTRISPEAPVPVVKVNTIEERPGAAVAMNI 60  
Db 1 MKVTLPEFERAGVWVGVDMLDRIYVGPTRISPEAPVPVVKVNTIEERPGAAVAMNI 60  
Qy 61 ASLGANARLVGLTGIDDAARALSKSLADVNCDFVSPHTITTKLRVLSRNOQLRLD 120  
Db 61 ASLGANARLVGLTGIDDAARALSKSLADVNCDFVSPHTITTKLRVLSRNOQLRLD 120  
Qy 121 FEEGEGVDPOPHERINQALSSIGALVLSDYAKGALASVOQMQLARKAGVPLIDPKG 180  
Db 121 FEEGEGVDPOPHERINQALSSIGALVLSDYAKGALASVOQMQLARKAGVPLIDPKG 180  
Qy 181 TDPERYRGATLTLPNLSFEFAVVGCKTBEIYERGMKLIADYELSLALVTRSEOGMSLL 240  
Db 181 TDPERYRGATLTLPNLSFEFAVVGCKTBEIYERGMKLIADYELSLALVTRSEOGMSLL 240  
Qy 241 QPGKAPLHPTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
Db 241 QPGKAPLHPTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
Qy 301 STVSPIELENAVGRADTGFVMTBEELKLAVAAARKGKGVMTNGVFDILAHGHSYL 360  
Db 301 STVSPIELENAVGRADTGFVMTBEELKLAVAAARKGKGVMTNGVFDILAHGHSYL 360  
Qy 361 ANARKDRLIVAVNSDASTKRLKDSRPVNPPLBORMIVLGALEAVDPVWVSFEEDTPORL 420  
Db 361 ANARKDRLIVAVNSDASTKRLKDSRPVNPPLBORMIVLGALEAVDPVWVSFEEDTPORL 420  
Qy 421 IAGILPDLVLKGGDYKPEEIASKEVWANGGEVLVNFEDGCSSTNNIKKIQQDKG 477

Db 421 IAGILPDLVLKGGDYKPEEIASKEVWANGGEVLVNFEDGCSSTNNIKKIQQDKG 477

RESULT 3  
US-09-815-242-10301  
Sequence 10301, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zykend, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10301  
LENGTH: 477  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-815-242-10301

Query Match 100.0%; Score 2393; DB 3; Length 477;  
Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MKVTLPEFERAGVWVGVDMLDRIYVGPTRISPEAPVPVVKVNTIEERPGAAVAMNI 60  
Db 1 MKVTLPEFERAGVWVGVDMLDRIYVGPTRISPEAPVPVVKVNTIEERPGAAVAMNI 60  
Qy 61 ASLGANARLVGLTGIDDAARALSKSLADVNCDFVSPHTITTKLRVLSRNOQLRLD 120  
Db 61 ASLGANARLVGLTGIDDAARALSKSLADVNCDFVSPHTITTKLRVLSRNOQLRLD 120  
Qy 121 FEEGEGVDPOPHERINQALSSIGALVLSDYAKGALASVOQMQLARKAGVPLIDPKG 180  
Db 121 FEEGEGVDPOPHERINQALSSIGALVLSDYAKGALASVOQMQLARKAGVPLIDPKG 180  
Qy 181 TDPERYRGATLTLPNLSFEFAVVGCKTBEIYERGMKLIADYELSLALVTRSEOGMSLL 240  
Db 181 TDPERYRGATLTLPNLSFEFAVVGCKTBEIYERGMKLIADYELSLALVTRSEOGMSLL 240  
Qy 241 QPGKAPLHPTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
Db 241 QPGKAPLHPTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
Qy 301 STVSPIELENAVGRADTGFVMTBEELKLAVAAARKGKGVMTNGVFDILAHGHSYL 360  
Db 301 STVSPIELENAVGRADTGFVMTBEELKLAVAAARKGKGVMTNGVFDILAHGHSYL 360  
Qy 361 ANARKDRLIVAVNSDASTKRLKDSRPVNPPLBORMIVLGALEAVDPVWVSFEEDTPORL 420

Db 361 ANARKDRLIVANVNSDASTKRLKGDSPVNPLEGRMIVGALAEVDMVVSFEEDTPORL 420  
Qy 421 IAGILPDLVKGSDYKPEBIAGSKYEWANGGEVLVNFEDGCSSTNIKKIQODKKG 477  
Db 421 IAGILPDLVKGSDYKPEBIAGSKYEWANGGEVLVNFEDGCSSTNIKKIQODKKG 477

## RESULT 4

US-10-282-122A-42614  
Sequence 42614, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Reminding Prior Application data removed - See file Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 42614  
LENGTH: 477  
TYPE: PRN  
ORGANISM: Escherichia coli  
US-10-282-122A-42614

Query Match 100.0%; Score 2393; DB 4; Length 477;  
Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKTTLPEFERAGVNVGDVMDLDRWYGPTRISPEAPVPPVKNTIIEERPGGAANYAMNI 60  
Db 1 MKTTLPEFERAGVNVGDVMDLDRWYGPTRISPEAPVPPVKNTIIEERPGGAANYAMNI 60  
Qy 61 ASIGANARLVGLGIDDAARALSKSLADVNVKCPVSPVPTHTTKRLVLSRNOQLIRLD 120  
Db 61 ASIGANARLVGLGIDDAARALSKSLADVNVKCPVSPVPTHTTKRLVLSRNOQLIRLD 120  
Qy 121 FEEGFEGVDPOPLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180  
Db 121 FEEGFEGVDPOPLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180

Qy 181 TDEERYGATLTPNISEPAVYKCKTEEBEYERGMKLIADYELSLVTRSEOGMSLL 240  
Db 181 TDEERYGATLTPNISEPAVYKCKTEEBEYERGMKLIADYELSLVTRSEOGMSLL 240  
Qy 241 QPKAPLHPPTQAEVYDVYAGDVIYGLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
Db 241 QPKAPLHPPTQAEVYDVYAGDVIYGLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
Qy 301 STVSPILENAVAGRADTGFGVMTBEBELKLAVAARRGKRVMTNGVFDILHAGVSYL 360  
Db 301 STVSPILENAVAGRADTGFGVMTBEBELKLAVAARRGKRVMTNGVFDILHAGVSYL 360  
Qy 361 ANARKDRLIVANVNSDASTKRLKGDSPVNPLEGRMIVGALAEVDMVVSFEEDTPORL 420  
Db 361 ANARKDRLIVANVNSDASTKRLKGDSPVNPLEGRMIVGALAEVDMVVSFEEDTPORL 420  
Qy 421 IAGILPDLVKGSDYKPEBIAGSKYEWANGGEVLVNFEDGCSSTNIKKIQODKKG 477  
Db 421 IAGILPDLVKGSDYKPEBIAGSKYEWANGGEVLVNFEDGCSSTNIKKIQODKKG 477

## RESULT 5

US-10-771-241-325  
Sequence 325, Application US/10771241  
Publication No. US20040241715A1  
GENERAL INFORMATION:  
APPLICANT: Zyskind, Judith  
APPLICANT: Forsyth, R. Allyn  
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
TITLE OF INVENTION: ESCHERICHIA COLI  
FILE REFERENCE: ELITRA.001C1  
CURRENT APPLICATION NUMBER: US/10/771,241  
CURRENT FILING DATE: 2004-02-03  
PRIOR APPLICATION NUMBER: 09/492,709  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/117,405  
PRIOR FILING DATE: 1999-01-27  
NUMBER OF SEQ ID NOS: 485  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 325  
LENGTH: 477  
TYPE: PRN  
ORGANISM: E. coli  
US-10-771-241-325

Query Match 100.0%; Score 2393; DB 5; Length 477;  
Best Local Similarity 100.0%; Pred. No. 1.3e-198;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKTTLPEFERAGVNVGDVMDLDRWYGPTRISPEAPVPPVKNTIIEERPGGAANYAMNI 60  
Db 1 MKTTLPEFERAGVNVGDVMDLDRWYGPTRISPEAPVPPVKNTIIEERPGGAANYAMNI 60  
Qy 61 ASIGANARLVGLGIDDAARALSKSLADVNVKCPVSPVPTHTTKRLVLSRNOQLIRLD 120  
Db 61 ASIGANARLVGLGIDDAARALSKSLADVNVKCPVSPVPTHTTKRLVLSRNOQLIRLD 120  
Qy 121 FEEGFEGVDPOPLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180  
Db 121 FEEGFEGVDPOPLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180  
Qy 181 TDEERYGATLTPNISEPAVYKCKTEEBEYERGMKLIADYELSLVTRSEOGMSLL 240  
Db 181 TDEERYGATLTPNISEPAVYKCKTEEBEYERGMKLIADYELSLVTRSEOGMSLL 240  
Qy 241 QPKAPLHPPTQAEVYDVYAGDVIYGLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
Db 241 QPKAPLHPPTQAEVYDVYAGDVIYGLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
Qy 301 STVSPILENAVAGRADTGFGVMTBEBELKLAVAARRGKRVMTNGVFDILHAGVSYL 360  
Db 301 STVSPILENAVAGRADTGFGVMTBEBELKLAVAARRGKRVMTNGVFDILHAGVSYL 360

```

Oy      361  ANARGGRLLVAVNSDSTKRLKQDSPVNFLEORMTVLGALBAVDWVVSFEEDTPORT 422
Db      361  ANARGGRLLVAVNSDSTKRLKQDSPVNFLEORMTVLGALBAVDWVVSFEEDTPORT 420
Oy      421  IAGILPDLVLVKGDDYKPEEIIAGSKKEVWANGSEVLVINFDGCGSTNNIIKKIQQDKG 477
Db      421  IAGILPDLVLVKGDDYKPEEIIAGSKKEVWANGSEVLVINFDGCGSTNNIIKKIQQDKG 477

RESULT 6
US-10-282-122A-56202
; Sequence 56202, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangyu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56202
; LENGTH: 476
; TYPE: prt
; ORGANISM: Enterobacter cloacae
US-10-282-122A-56202

```

Query Match	95.9%	Score 2294	DB 4:	Length 476
Best Local Similarity	94.7%	Pred. No. 5.1e-190		
Matches 451	Conservative 19	Matches 6	Indels 0	Gaps 0
Qy	1	MKVTLPPEERAGVMVGVMDLDRMYGPFPSRISPEAPRVVVKVNTIEERPGGANVANI	60	
Db	1	MKVTLPPEERAGVMVGVMDLDRMYGPFPSRISPEAPRVVVKVNTIEERPGGANVANI	60	
Qy	61	ASLGNARLVGLTGIDDAARALSKSLADVNVKDFVSVPETHPTITKLRLVSRNQILRLD	120	
Db	61	ASLGNARLVGLTGIDDAARALSKSLADVNVKDFVSVPETHPTITKLRLVSRNQILRLD	120	
Qy	121	FEBSGEGVDPDLPIHERINQALSSIGALVLSDYAKGALASVQOMTQLARRAGVPVLTIDPKG	180	

Db	121	FEBSGEGVDPBEBLHRIQALGNIALVSDYAKGALLASVKMIIQLARKANPVILDPKQ	180
QY	181	TDPERYRGATILITPNLSFEFAVVGCKTEEBEIVBERGMKLIADYELSALLVTRSEQMSLL	240
Db	181	TDPERYRGATILITPNLSFEFAVVGCKTEEBEIVBERGMKLIADYELSALLVTRSEQMSLL	240
QY	241	QPGKRLPMPPOAGVYDVYTAGADPVIVUAAITLAAGNSLEBACFPANAAAGVYVKELGT	300
Db	241	QPGKRLPMPPOAGVYDVYTAGADPVIVUAAITLAAGNSLEBACFPANAAAGVYVKELGT	300
QY	301	STVSEIIELENNVVRGADYTGFGVMTSEELKLAANAARKGEEKVMTNGVPEDIILAHGVSYL	360
Db	301	STVSEIIELENNVVRGADYTGFGVMTSEELKLAANAARKGEEKVMTNGVPEDIILAHGVSYL	360
QY	361	ANARKLGRLIVAANSDSATYRLKCDSPVNDLEOPMIVLGALEAVDVMVVSFEEDTPORL	420
Db	361	ANARKLGRLIVAANSDSATYRLKCDSPVNDLEOPMIVLGALEAVDVMVVSFEEDTPORL	420
QY	421	IAGILPDLLVVGSGDYKPEEIRIAGSEKVMVANGCVLVLINPEDEGSTNNIIKKIQODKK	476
Db	421	IAGILPDLLVVGSGDYKPEEIRIAGSEKVMVANGCVLVLINPEDEGSTNNIIKKIQODSQ	476

```

RESULT 7
; US-09-815-242-13759
; Sequence 13759, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13759
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Salmonella typhi
; US-09-815-242-13759

```

	Query Match	94.7%	Score 2265	DB 3	Length 477
	Best Local Similarity	93.5%	Pred. No.	1.7e-187	
	Matches	445	Conservative	19	Mismatches 12; Indels 0; Gaps 0
OY	1 MKVTVLPEEPKAGVMVGVMTDRYYKKGPSTRISPEAPVVVVVVNTTEERPGGAANAMNI	60			
Db	1 MKNVTLPAEPRGAVMVGGVMDLRITVGPCRSPEAPVVVVKNATVEERPGGAANAMNI	60			
OY	61 ASLGNAARLVGLTIGIDDAARALSKSLADVNVCDFSVETHTPTITLYRVLSSNQQLRLD	120			

Db 61 ASIGANRLVGLTGTIDDAARLSKTLAEVNVKDFVSPHPTITTKLRVSRNOQLRLD 120  
Qy 121 FEESGEGVDPPQLHERINQALSSIGALVLSDYAKGALASVOQMIOLARKAGVPLIDPKG 180  
Db 121 FEESGEGVDPPQLHERINQALSSIGALVLSDYAKGALASVOQMIOLARKAGVPLIDPKG 180  
Qy 161 TDFERYGATLLTPNISEFEAVVCKCTEERIVERGKLIADYELSLALVTRSEOGMTLL 240  
Db 181 TDFERYGATLLTPNISEFEAVVCKCTEERIVERGKLIADYELSLALVTRSEOGMTLL 240  
Qy 241 QPKAPLHPMTQAOEYVDYTGADTVIGVLAATLAAGNSLEACFPANAAAGVVGGLGT 300  
Db 241 QPKAPLHPMTQAOEYVDYTGADTVIGVLAATLAAGNSLEACFPANAAAGVVGGLGT 300  
Qy 301 STVSPIELBNVGRADTGFVMTSEELKLAVAARRGKGVMTNGVPDILHAGVSYL 360  
Db 301 STVSPIELBNVGRADTGFVMTSEELKLAVAARRGKGVMTNGVPDILHAGVSYL 360  
Qy 361 ANARKLGDRLIVANSASDASTKRLKGSRPVNPLEQRMIVGALSVDMVVSFEEDTPORL 420  
Db 361 ANARKLGDRLIVANSASDASTKRLKGSRPVNPLEQRMIVGALSVDMVVSFEEDTPORL 420  
Qy 421 IAGILPDLVKGDDYKPEEINGSKRVANGGEVLVNFEDCGSTTNIIKKIQODES 476  
Db 421 IAGILPDLVKGDDYKPEEINGSKRVANGGEVLVNFEDCGSTTNIIKKIQODES 476

RESULT 8  
US-10-282-122A-75478  
Sequence 75478, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282.122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 75478

LENGTH: 477  
TYPE: PRF  
ORGANISM: Salmoneilla typhi  
US-10-282-122A-75478  
Query Match 94.7%; Score 2265; DB 4; Length 477;  
Best Local Similarity 93.5%; Pred. No. 1,7e-187;  
Matches 445; Conservative 19; Mismatches 12; Indels 0; Gaps 0;  
Qy 1 MKYTLPEFERAGVNVGVDMKLDNRYTGFTSRISPEAPVYKNTTIEERGGANVAMNI 60  
Db 1 MKYTLPEFERAGVNVGVDMKLDNRYTGFTSRISPEAPVYKNTTIEERGGANVAMNI 60  
Qy 61 ASIGANRLVGLTGTIDDAARLSKTLAEVNVKDFVSPHPTITTKLRVSRNOQLRLD 120  
Db 61 ASIGANRLVGLTGTIDDAARLSKTLAEVNVKDFVSPHPTITTKLRVSRNOQLRLD 120  
Qy 121 FEESGEGVDPPQLHERINQALSSIGALVLSDYAKGALASVOQMIOLARKAGVPLIDPKG 180  
Db 121 FEESGEGVDPPQLHERINQALSSIGALVLSDYAKGALASVOQMIOLARKAGVPLIDPKG 180  
Qy 181 TDFERYGATLLTPNISEFEAVVCKCTEERIVERGKLIADYELSLALVTRSEOGMTLL 240  
Db 181 TDFERYGATLLTPNISEFEAVVCKCTEERIVERGKLIADYELSLALVTRSEOGMTLL 240  
Qy 241 QPKAPLHPMTQAOEYVDYTGADTVIGVLAATLAAGNSLEACFPANAAAGVVGGLGT 300  
Db 241 QPKAPLHPMTQAOEYVDYTGADTVIGVLAATLAAGNSLEACFPANAAAGVVGGLGT 300  
Qy 301 STVSPIELBNVGRADTGFVMTSEELKLAVAARRGKGVMTNGVPDILHAGVSYL 360  
Db 301 STVSPIELBNVGRADTGFVMTSEELKLAVAARRGKGVMTNGVPDILHAGVSYL 360  
Qy 361 ANARKLGDRLIVANSASDASTKRLKGSRPVNPLEQRMIVGALSVDMVVSFEEDTPORL 420  
Db 361 ANARKLGDRLIVANSASDASTKRLKGSRPVNPLEQRMIVGALSVDMVVSFEEDTPORL 420  
Qy 421 IAGILPDLVKGDDYKPEEINGSKRVANGGEVLVNFEDCGSTTNIIKKIQODES 476  
Db 421 IAGILPDLVKGDDYKPEEINGSKRVANGGEVLVNFEDCGSTTNIIKKIQODES 476

RESULT 9  
US-10-282-122A-60277  
Sequence 60277, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282.122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09



PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 60277  
LENGTH: 477  
TYPE: PRN  
ORGANISM: Klebsiella pneumoniae  
US-10-282-122A-60277

Query Match 93.1%; Score 2227; DB 4; Length 477;  
Best Local Similarity 92.2%; Pred. No. 3,3e-184;  
Matches 437; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

QY 1 MKVTLPEERAGVNVGDVMDRKYGPTSRISPEAPVYKNTIERRPGGANVAMNI 60  
DB 1 MKVTLPEERAGVNVGDVMDRKYGPTSRISPEAPVYKNTIERRPGGANVAMNI 60  
QY 61 ASLGANRLVGLTGIDDAARALSKSLADVNCDFVSPHTPTTKLRVLSRNOQLRLD 120  
DB 61 ASLGANRLVGLTGIDDAARALSKSLADVNCDFVSPHTPTTKLRVLSRNOQLRLD 120  
QY 121 FEBGFEVDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180  
DB 121 FEBGFEVDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180  
QY 121 FEBGFEVDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180  
DB 121 FEBGFEVDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKG 180  
QY 181 TDFERYKATLTTPMLSEFAVNGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240  
DB 181 TDFERYKATLTTPMLSEFAVNGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240  
QY 181 TDFERYKATLTTPMLSEFAVNGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240  
DB 181 TDFERYKATLTTPMLSEFAVNGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240  
QY 241 QGKAPLPMPTQAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLT 300  
DB 241 QGKAPLPMPTQAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLT 300  
QY 241 QGKAPLPMPTQAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLT 300  
DB 241 QGKAPLPMPTQAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLT 300  
QY 301 STVPSELENAVRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYL 360  
DB 301 STVPSELENAVRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYL 360  
QY 301 STVPSELENAVRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYL 360  
DB 301 STVPSELENAVRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYL 360  
QY 361 ANARKLGRLLVAVNSDASTKRLKDSRPVNPLEQRMIVLGALBAVDMVVSFEEDTQRL 420  
DB 361 ANARKLGRLLVAVNSDASTKRLKDSRPVNPLEQRMIVLGALBAVDMVVSFEEDTQRL 420  
QY 421 IAGILPDLVKGQDYKPEEIASGEVWANGGEVLVNFEDGCSITTNIKKIQ 474  
DB 421 IAGILPDLVKGQDYKPEEIASGEVWANGGEVLVNFEDGCSITTNIKKIQ 474

RESULT 10

US-10-282-122A-73168  
Sequence 73168, Application US/10282122A  
Publication NO. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Hagelbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: EUTRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 73168  
LENGTH: 463  
TYPE: PRN  
ORGANISM: Salmonella paratyphi A  
US-10-282-122A-73168

Query Match 91.8%; Score 2196; DB 4; Length 463;  
Best Local Similarity 93.5%; Pred. No. 1.6e-181;  
Matches 435; Conservative 18; Mismatches 10; Indels 2; Gaps 2;

QY 8 FEBAGVNVGDVMDRKYGPTSRISPEAPVYKNTIERRPGGANVAMNIASIGANA 67  
DB 1 FEBAGVNVGDVMDRKYGPTSRISPEAPVYKNTIERRPGGANVAMNIASIGANA 67  
QY 68 RLVGLTGIDDAARALSKSLADVNCDFVSPHTPTTKLRVLSRNOQLRLDPEBEG 127  
DB 68 RLVGLTGIDDAARALSKSLADVNCDFVSPHTPTTKLRVLSRNOQLRLDPEBEG 127  
QY 61 RLVGLTGIDDAARALSKSLADVNCDFVSPHTPTTKLRVLSRNOQLRLDPEBEG 119  
DB 61 RLVGLTGIDDAARALSKSLADVNCDFVSPHTPTTKLRVLSRNOQLRLDPEBEG 119  
QY 128 VDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKGTFERYR 187  
DB 128 VDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKGTFERYR 187  
QY 120 VDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKGTFERYR 179  
DB 120 VDPQLHERINQALSSIGALVLSDYAKGALASVOQMIQLARKAGVPLIDPKGTFERYR 179  
QY 188 GATLTPNLSEFAVNGKCTEERIVERGKMLADYELSLALVTRSEQMSLLQPKAPL 247  
DB 188 GATLTPNLSEFAVNGKCTEERIVERGKMLADYELSLALVTRSEQMSLLQPKAPL 247  
QY 248 HMTPOAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLTSTVPSE 307  
DB 248 HMTPOAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLTSTVPSE 307  
QY 240 HMTPOAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLTSTVPSE 299  
DB 240 HMTPOAOEVYVNTGAGPTVIGVLAATLAAGNSLEACFPANAAAGVVGKLTSTVPSE 299  
QY 308 LENA VRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYLANARKLG 367  
DB 308 LENA VRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYLANARKLG 367  
QY 300 LENA VRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYLANARKLG 358  
DB 300 LENA VRGADTGFVMTTEEBELKLAVAARKEGKVYNTGVFDILAHGHSYLANARKLG 358  
QY 368 DRLLIVAVNSDASTKRLKDSRPVNPLEQRMIVLGALBAVDMVVSFEEDTQRLIAGILPD 427  
DB 368 DRLLIVAVNSDASTKRLKDSRPVNPLEQRMIVLGALBAVDMVVSFEEDTQRLIAGILPD 427  
QY 359 DRLLIVAVNSDASTKRLKDSRPVNPLEQRMIVLGALBAVDMVVSFEEDTQRLIAGILPD 418  
DB 359 DRLLIVAVNSDASTKRLKDSRPVNPLEQRMIVLGALBAVDMVVSFEEDTQRLIAGILPD 418  
QY 428 LTVKGDYKPEEIASGEVWANGGEVLVNFEDGCSITTNIKKIQ 472  
DB 428 LTVKGDYKPEEIASGEVWANGGEVLVNFEDGCSITTNIKKIQ 472  
QY 419 LTVKGDYKPEEIASGEVWANGGEVLVNFEDGCSITTNIKKIQ 463  
DB 419 LTVKGDYKPEEIASGEVWANGGEVLVNFEDGCSITTNIKKIQ 463

RESULT 11

US-10-282-122A-78165  
Sequence 78165, Application US/10282122A

```
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyckind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 78165
LENGTH: 476
TYPE: PRT
ORGANISM: Yersinia pestis
US-10-282-122A-78165

Query Match      86.8%; Score 2076; DB 4; Length 476;
Best Local Similarity 86.1%; Pred. No. 4,2e-171;
Matches 409; Conservative 32; Mismatches 34; Indels 0; Gaps 0;

QY 1 MKVTLPEFRAGVNVGVDMLDRIYWGPTSRISPEAPVPVVKVNTIEERPGAAVAMNI 60
DB 1 MKVTLPEFRAGVNVGVDMLDRIYWGPTSRISPEAPVPVVKVNTIEERPGAAVAMNI 60
QY 61 ASIGANRLVGLTGIDDAARALSKSLADVNVKCDPVSVPHPTTKRLVLSRNOQLRLD 120
DB 61 ASIGANRLVGLTGIDDAARALSKSLADVNVKCDPVSVPHPTTKRLVLSRNOQLRLD 120
QY 121 FEESFEVDVDFPFLERINQALSLISIGALVLSDYAGALASVOOMIQLARKAGVPYLIDPKG 180
DB 121 FEESFEVDVDFPFLERINQALSLISIGALVLSDYAGALASVOOMIQLARKAGVPYLIDPKG 180
QY 181 TDFERYGATLTLLTNLSEFEAVGKCTEERIVERGKMLADYELSLALLVTRSEQMSLL 240
DB 181 TDFERYGATLTLLTNLSEFEAVGKCTEERIVERGKMLADYELSLALLVTRSEQMSLL 240
QY 241 QPGKAPLHMPQAOBVDVNGAGDPTVGVLAATLAAGNSLEAEPFANAAAGVVGKLGCT 300
DB 241 QPGKAPLHMPQAOBVDVNGAGDPTVGVLAATLAAGNSLEAEPFANAAAGVVGKLGCT 300
QY 301 STVSPLEENAVNRADDTGFGVWTEEBELKLAVAARKGKRVWNTNGVFDTIILHAGVSYL 360
DB 301 STVSPLEENAVNRADDTGFGVWTEEBELKLAVAARKGKRVWNTNGVFDTIILHAGVSYL 360
```

```
DB 301 STVSPLEENAVNRGRATGFGVWDEBOQLKIAVAQARORGEKRVWNTNGIFDIIHAGVSYL 360
QY 361 ANARKGDRILVAVNSDASKRLKGSRPVNPLEOMITVIGALEAVDMVWVFEEDTPORL 420
DB 361 ANARKGDRILVAVNSDASKRLKGSRPVNPLEOMITVIGALEAVDMVWVFEEDTPORL 420
QY 421 IAGILPDLVKGADYKPBBIASKEVWANGSEVLYNFEDGCTNIIKKIQODK 475
DB 421 IAGILPDLVKGADYKPBBIASKEVWANGSEVLYNFEDGCTNIIKKIQODK 475

RESULT 12
US-10-282-122A-68720
Sequence 68720, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyckind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 68720
LENGTH: 474
TYPE: PRT
ORGANISM: Proteus mirabilis
US-10-282-122A-68720

Query Match      83.9%; Score 2007; DB 4; Length 474;
Best Local Similarity 82.2%; Pred. No. 4e-165;
Matches 389; Conservative 43; Mismatches 41; Indels 0; Gaps 0;

QY 1 MKVTLPEFRAGVNVGVDMLDRIYWGPTSRISPEAPVPVVKVNTIEERPGAAVAMNI 60
DB 1 MKVTLPEFRAGVNVGVDMLDRIYWGPTSRISPEAPVPVVKVNTIEERPGAAVAMNI 60
QY 61 ASIGANRLVGLTGIDDAARALSKSLADVNVKCDPVSVPHPTTKRLVLSRNOQLRLD 120
DB 61 ASIGANRLVGLTGIDDAARALSKSLADVNVKCDPVSVPHPTTKRLVLSRNOQLRLD 120
```

Qy 121 FEEBEGVDPOPLHERINQALSTGALVLSYANGALASVOQMTQLARKGVPVLIDPKG 180  
Db 121 FEEBFSNVSPPIYERIQOALPSWGALILSDYAKGALSHVOQMTQLAKKGVPLIDPKG 180  
Qy 181 TDFERYGATLLTNTLSFEELVNGKCTEERIVERGKMLIDYELSLVLTSSQGSML 240  
Db 181 SDFERYKATLLTPNMSFEBOVGCHTDDLVKRGQVLDLADLLITRSRGSML 240  
Qy 241 QPGKAPLHPTQAOEYVDVTGAGDTVIGVLAATLAAGNSLEBACFPANAAGVVGKLGST 300  
Db 241 QANAPLHPTQAOEYVDVTGAGDTVIGVLAATLAAGKPLAEACFLANAAAGVVGKLGST 300  
Qy 301 STVPPIELHNAVRGADTGFVMTBEEKLAVAARKEGKVMTNGVPIIHAHVSYL 360  
Db 301 STVPPIELHNAVRGADTGFVMTBEEKLAVAARKEGKVMTNGVPIIHAHVSYL 360  
Qy 361 ANARKLGDRLIVANSNDASTKRLKGDSPVNPLEORMTVLGALAVDMVVSFEEDTPORL 420  
Db 361 ANARKLGDRLIVANSNDASTKRLKGDSPVNPLEORMTVLGALAVDMVVSFEEDTPORL 420  
Qy 421 IAGILPDLIVKGDYKPEEIASKEVWANGGEVLVLPEDGCTTNIKKIQ 473  
Db 421 IASVLPDLIVKGDYKPEEIASKEVWANGGEVLVLPEDGCTTNIKKIK 473

RESULT 13  
US-10-282-122A-58533  
Sequence 58533, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT FILING DATE: 2003-02-20  
CURRENT APPLICATION NUMBER: US/10/282,122A  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See file wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: Patent version 3.1  
SEQ ID NO 58533  
LENGTH: 476  
TYPE: PRT  
ORGANISM: Haemophilus influenzae

US-10-282-122A-58533  
Query Match 71.5%; Score 1712; DB 4; Length 476;  
Best Local Similarity 71.0%; Pred. No. 1,6e-139;  
Matches 331; Conservative 58; Mismatches 77; Indels 0; Gaps 0;  
Qy 7 EFERAGVNVGVMDRMYGPTSRISPEAPVNVKNTITEEPGGAANYANNIISIGAN 66  
Db 7 EFRQAKVLVLDGVMILRMYFGATNRISPEAPVNVVQENEERAGGAANYANNIISINVP 66  
Qy 67 ARVGLTGIDDAARLSKSLADVNVKDFVSVTHPTITGLRVLSNQOILRLDPEEFGFE 126  
Db 67 VQLMGLIGDDETSALSLIBEKOKIDCNFVALTHTHTTKRLISRHQILRLDPEEDFN 126  
Qy 127 GVDPOPLHERINQALSTGALVLSYANGALASVOQMTQLARKGVPVLIDPKGTDFERY 186  
Db 127 NVDCOKLLALASAVNMGALILSDYKGTLLKOVQMTQLARKANVPVLIDPKGTDFERY 186  
Qy 187 RGATLLTPNISFEPAVNGKCTEERIVERGKMLIDYELSLVLTSSQGSMLDQKAP 246  
Db 187 RGATLLTPNMSFEPAVNGKCTEERIVERGKMLISDIELTALLVTRSEKMTLLRPQEP 246  
Qy 247 LHMPTQAOEYVDVTGAGDTVIGVLAATLAAGNSLEBACFPANAAGVVGKLGSTVSP 306  
Db 247 YHLPVAKGVFDVTGAGDTVIGVLAATLAADGNSFEESCYNANAGIVGKLGSTVSTV 306  
Qy 307 ELENAVGRADTGFVMTBEEKLAVAARKEGKVMTNGVPIIHAHVSYLANARKL 366  
Db 307 ELENAIHAREPTFGIMSEBELDVAQAKAREKLVMTNGCGDILHPHISILENARKL 366  
Qy 367 GDRILIVANSNDASTKRLKGDSPVNPLEORMTVLGALAVDMVVSFEEDTPORLIAGILP 426  
Db 367 GDRILIVANSDDSVKRLKGDSPVNPINNIENRMAVLGASVDMVLPPTEDPQRLIGILP 426  
Qy 427 DILVKGADYKPEEIASKEVWANGGEVLVLPEDGCTTNIKKIQ 472  
Db 427 DILVKGADYKPEEIASKEVWANGGEVLVLPEDGCTTNIKKIK 472

RESULT 14  
US-10-282-122A-67124  
Sequence 67124, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT FILING DATE: 2003-02-20  
CURRENT APPLICATION NUMBER: US/10/282,122A  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 67124  
 LENGTH: 476  
 TYPE: PRT  
 ORGANISM: Pasteurella multocida  
 US-10-282-122A-67124

Query Match 71.1%; Score 1701; DB 4; Length 476;  
 Best Local Similarity 69.7%; Pred. No. 1.5e-138;  
 Matches 327; Conservative 65; Mismatches 77; Indels 0; Gaps 0;

7 EFEEAGVNVGVNDVMDRYYGPTSRISPEAPVPPVKNVTEERPGGANVAMNTASIGAN 66  
 7 EFKAQKLVLDVMDRYYGPTSRISPEAPVPPVKNVTEERPGGANVAMNTASIGAN 66  
 67 ARLVGLGIDDAARALSKSLADVNVKDFVSPHPTTKLRVLSRNOQLIRLDFEESGE 126  
 67 VKLGLTGRDERTGLAESLAKSRIECDPVOJLHPTTKLRISRNOQLIRLDFEEDFQ 126  
 127 GVDDPQHERINQALSSIGALVLSDYAKGALASVOQMIOLARKAGVPLIDPKGTDFERY 186  
 127 NVTSDALLOKLDALQNVGALVLSDYKGKTLNVOQMIOLARKAGVPLIDPKGTDFERY 186  
 187 RGATLTPNISEFAVNVKCTEERIEVERGKLIADYELSLVTRSEQGSLLQPGKAP 246  
 187 RGATLTPNISEFAVNVKCTEERIEVERGKLIADYELSLVTRSEQGSLLQPGKAP 246  
 247 LHMPTQAOEYVDVAGDPTVIGVLAATLAAGNSLEBACFPANAAAGVVGGLGTSTVSP 306  
 247 FHLTEKEVVDVAGDPTVIGVLAATLAAGNSLEBACFPANAAAGVVGGLGTSTVSP 306  
 307 ELENAVGRADTGFVMTBEELKLAVAARRGKRVMTNGVFDILHAGHVSYLENARKL 366  
 307 ELENAVGRADTGFVMTBEELKLAVAARRGKRVMTNGVFDILHAGHVSYLENARKL 366  
 367 GDRILIVANSDASVRLKSGSRPNPPEORNTIVGALBAVDVVSFEEDTPQRLIAGTLP 426  
 367 GDRILIVANSDASVRLKSGSRPNPPEORNTIVGALBAVDVVSFEEDTPQRLIAGTLP 426  
 427 DLVKGQDYKPEEIASGKEVWANGBEVLVLFEDGCTTNIKKIQODK 475  
 427 DLVKGQDYKPEEIASGKEVWANGBEVLVLFEDGCTTNIKKIQODK 475

RESULT 15  
 US-10-282-122A-69312  
 Sequence 69312, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Karl  
 APPLICANT: Zykkind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Foraych, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: EUTRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/220,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 69312  
 LENGTH: 474  
 TYPE: PRT  
 ORGANISM: Pseudomonas syringae  
 US-10-282-122A-69312

Query Match 56.7%; Score 1357.5; DB 4; Length 474;  
 Best Local Similarity 58.1%; Pred. No. 9.3e-109;  
 Matches 274; Conservative 70; Mismatches 127; Indels 1; Gaps 1;

1 MKYTLPEFERAGVNVGVNDVMDRYYGPTSRISPEAPVPPVKNVTEERPGGANVAMNT 60  
 1 MKLSMPRFQDAPLVVGVNDVMDRYYGPTSRISPEAPVPPVKNVTEERPGGANVAMNT 60  
 61 ASIGANARLVGLGIDDAARALSKSLADVNVKDFVSPHPTTKLRVLSRNOQLIRL 120  
 61 AALGAPASLVGVGDEMAASLTNSLPAAGVVARFORIADQPTTVLRVNSRNOQLIRL 120  
 121 FEESGEGVDPPQHERINQALSSIGALVLSDYAKGALASVOQMIOLARKAGVPLIDPKG 180  
 121 FEESP-NTDPLASAVYSLIDGKTVLSDYKGKALKHQALIOARKRGIVLADPKG 179  
 181 TDPERYRGATLTPNISEFAVNVKCTEERIEVERGKLIADYELSLVTRSEQGSLL 240  
 180 KDFAIRGASLTPNISEFAVNVGHCEDEAQLVTKGALQMLQEBDLGLVTRGEHGTLL 239  
 241 QPGKAPLHMPTQAOEYVDVAGDPTVIGVLAATLAAGNSLEBACFPANAAAGVVGGLGT 300  
 240 RPDQALHLPARAREVDVAGDPTVISTLAAAIAAGEELPHAVALANLAAIVGGLGT 299  
 301 STVSPTELENAVGRADTGFVMTBEELKLAVAARRGKRVMTNGVFDILHAGHVSYLE 360  
 300 AALISAPELRPAITOREBSESGVGLGELLLAIVDADARHKEKIPFNGCFDILHAGHVTYL 359  
 361 ANARKLGDRLIVANSDASVRLKSGSRPNPPEORNTIVGALBAVDVVSFEEDTPQRL 420  
 360 EQARAGDRLIVANSDASVRLKSGSRPNPINSVDRBAVLAGAGVDMVVSFEEDTPQRL 419  
 421 IAGILPVLVKGQDYKPEEIASGKEVWANGBEVLVLFEDGCTTNIKKIQ 472  
 420 LTHVKPDVLVKGQDYGIDGVVAGDIYQAVYGBVRLVGLVENSSTIAIVEKIR 471

Search completed: March 18, 2006, 06:21:19  
 Job time: 170 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model1

Run on: March 19, 2006, 11:24:39 ; Search time 2093 Seconds

(without alignments)  
5665.689 Million cell updates/sec

Title: US-10-520-820-28

Perfect score: 1434

Sequence: 1 atgaaagtaacgctgcgcaga.....aacagataaaagctaa 1434

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1434	100.0	US-09-741-669-276	Sequence 276, App
2	1434	100.0	US-09-912-020-165	Sequence 165, App
3	1434	100.0	US-09-815-242-6204	Sequence 6204, Ap
4	1434	100.0	US-10-282-122A-6430	Sequence 6430, Ap
5	1434	100.0	US-10-771-241-165	Sequence 165, App
6	1018.2	71.0	US-10-282-122A-20018	Sequence 20018, A
7	1008.6	70.3	US-09-815-242-6661	Sequence 9661, Ap
8	1008.6	70.3	US-10-282-122A-39294	Sequence 39294, A
9	1007	70.2	US-10-282-122A-24093	Sequence 36984, A
10	935.2	65.3	US-10-282-122A-41981	Sequence 41981, A
11	879.2	61.3	US-10-282-122A-32536	Sequence 32536, A
12	725	50.6	US-10-282-122A-22349	Sequence 22349, A
13	635.2	44.3	US-10-329-670-1	Sequence 1, App11
14	635.2	44.3	US-10-158-865-1	Sequence 1, App11
15	635.2	44.3	US-10-981-687-1	Sequence 1, App11
16	635.2	44.3	US-10-795-159-685	Sequence 685, App
17	625.6	43.6	US-10-282-122A-30940	Sequence 30940, A
18	617.4	43.1	US-10-363-345A-7681	Sequence 7681, Ap
19	507.8	35.4	US-10-363-345A-7682	Sequence 7682, Ap
20	507.8	35.4	US-10-363-345A-7681	Sequence 7681, Ap
21	507.8	35.4	US-10-363-345A-7682	Sequence 7682, Ap
22	497.8	34.7	US-09-815-242-7981	Sequence 7981, Ap

24	497.8	34.7	1425	7	US-10-282-122A-30585	Sequence 30585, A
25	493	34.4	1422	7	US-10-282-122A-33128	Sequence 33128, A
26	465.6	32.5	1407	8	US-10-282-122A-31665	Sequence 31665, A
27	425.6	29.7	707	8	US-10-363-345A-7683	Sequence 7683, Ap
28	425.2	29.7	707	8	US-10-363-345A-7684	Sequence 7684, Ap
29	425.2	29.7	707	9	US-10-363-345A-7684	Sequence 7684, Ap
30	425.2	29.7	707	9	US-10-363-345A-7684	Sequence 7684, Ap
31	406.2	28.3	2055	9	US-10-450-763-25673	Sequence 25673, A
32	397	27.7	555	8	US-10-363-345A-32035	Sequence 32035, A
33	397	27.7	555	8	US-10-363-345A-32036	Sequence 32036, A
34	397	27.7	555	8	US-10-363-345A-32036	Sequence 32036, A
35	397	27.7	555	9	US-10-363-345A-32036	Sequence 32036, A
36	390.4	27.2	1001	7	US-10-343-561-64	Sequence 64, App1
37	388.6	27.1	3641	9	US-10-795-159-458	Sequence 458, App
38	340.8	23.8	555	8	US-10-363-345A-32033	Sequence 32033, A
39	340.8	23.8	555	8	US-10-363-345A-32034	Sequence 32034, A
40	340.8	23.8	555	9	US-10-363-345A-32033	Sequence 32033, A
41	340.8	23.8	555	9	US-10-363-345A-32034	Sequence 32034, A
42	302.6	21.1	706	8	US-10-363-345A-7297	Sequence 7297, Ap
43	302.6	21.1	706	8	US-10-363-345A-7298	Sequence 7298, Ap
44	302.6	21.1	706	9	US-10-363-345A-7297	Sequence 7297, Ap
45	302.6	21.1	706	9	US-10-363-345A-7298	Sequence 7298, Ap

## ALIGNMENTS

RESULT 1  
US-09-741-669-276  
Sequence 276, Application US/09741669  
Patent No. US200202718A1  
GENERAL INFORMATION:  
APPLICANT: Forbyth, R. Allyn  
APPLICANT: Ohlsen, Karl L.  
TITLE OF INVENTION: Zykkind, Judith W.  
TITLE OF INVENTION: Genes identified as required for  
FILE REFERENCE: ELITRA, 009A  
CURRENT APPLICATION NUMBER: US/09/741.669  
CURRENT FILING DATE: 2000-12-19  
PRIOR APPLICATION NUMBER: US 60/173005  
NUMBER OF SEQ ID NOS: 481  
SOFTWARE: PaastSeq for Windows Version 4.0  
SEQ ID NO 276  
LENGTH: 1434  
TYPE: DNA  
ORGANISM: Escherichia coli  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1434)  
US-09-741-669-276  
Query Match 100.0%; Score 1434; DB 3; Length 1434;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAGTAACGCTGCCAGAGTTGAACGTGCAGAGTATGTTGGTATGATG 60  
DB 1 ATGAAAGTAACGCTGCCAGAGTTGAACGTGCAGAGTATGTTGGTATGATG 60  
QY 61 CTGATCTCTACCTGATGACGCGCCACAGTCGATCTCGCGGAAGCGCGGCGCG 120  
DB 61 CTGATCTCTACCTGATGACGCGCCACAGTCGATCTCGCGGAAGCGCGGCGCG 120  
QY 121 GTTAAAGTAATACATTCAGAGAGCTCCGCGCGCGCTAATCGTGCATGATATC 180  
DB 121 GTTAAAGTAATACATTCAGAGAGCTCCGCGCGCGCTAATCGTGCATGATATC 180  
QY 181 GTTCTCTCTGCTCTAATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 181 GTTCTCTCTGCTCTAATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

QY	241	GC	CGCTAGTAAATCTCTGCGCGAGCTGAACGTCAAAAGCGACTTGTTCTGTACCGAAG	300
Db	241	GC	CGCTAGTAAATCTCTGCGCGAGCTGAACGTCAAAAGCGACTTGTTCTGTACCGAAG	300
QY	301	CAT	CGACCACTTACCAATTAACGGGTATCTTCCGCAACCAACAGCTGATCCGCTGAT	360
Db	301	CAT	CGACCACTTACCAATTAACGGGTATCTTCCGCAACCAACAGCTGATCCGCTGAT	360
QY	361	TTT	GAAAGAGTTTCGAAAGTGTTGATCCGAGCCGCTGCACGAGCGATTAATCAGCG	420
Db	361	TTT	GAAAGAGTTTCGAAAGTGTTGATCCGAGCCGCTGCACGAGCGATTAATCAGCG	420
QY	421	CTG	AGTTGATTTGGGCGCTGCTGTTCTCTACATACGCGCAAAAGGTGCGCTGGAAAGGTA	480
Db	421	CTG	AGTTGATTTGGGCGCTGCTGTTCTCTACATACGCGCAAAAGGTGCGCTGGAAAGGTA	480
QY	481	CAG	CAGATGATCCAACTGCGCGGTAAACGGGTGTTCCGCTGCTGATTTGATCCAAAGGT	540
Db	481	CAG	CAGATGATCCAACTGCGCGGTAAACGGGTGTTCCGCTGCTGATTTGATCCAAAGGT	540
QY	541	ACCG	ATTTTGAAGCCGCTACCGCGCGCTACGCTGTTAAGCGCAATCTCTCGAAATTTGAA	600
Db	541	ACCG	ATTTTGAAGCCGCTACCGCGCGCTACGCTGTTAAGCGCAATCTCTCGAAATTTGAA	600
QY	601	GCT	GTTGTCGGTAAATGTTAAGACCAAGAAAGATTTGTAGCCGCGCATAACTGATT	660
Db	601	GCT	GTTGTTGCGGTAAATGTTAAGACCAAGAAAGATTTGTAGCCGCGCATAACTGATT	660
QY	661	GCC	GATTAAGAACTCTCGGCTCTGTGTAAGACCGTTCGAAACAGGGTATGCTGCTG	720
Db	661	GCC	GATTAAGAACTCTCGGCTCTGTGTAAGACCGTTCGAAACAGGGTATGCTGCTG	720
QY	721	CA	ACCGGTTAAAGCGCGCTGATATGCGAACCCAGCGCAGGAAGTATGACGTTAC	780
Db	721	CA	ACCGGTTAAAGCGCGCTGATATGCGAACCCAGCGCAGGAAGTATGACGTTAC	780
QY	781	GGT	GCGCGCGACACGCTGATTTGGCTCTTGCGCGCAACGCTGCGACGCGGTAAATTCG	840
Db	781	GGT	GCGCGCGCGACACGCTGATTTGGCTCTTGCGCGCAACGCTGCGACGCGGTAAATTCG	840
QY	841	GAA	GAAAGCCTGCTTTCTTGCCAAATCGCGCGCTGCGTGGTGTGCGCAAACTGGGAAC	900
Db	841	GAA	GAAAGCCTGCTTTCTTGCCAAATCGCGCGCTGCGTGGTGTGCGCAAACTGGGAAC	900
QY	901	TC	CACGGTTTCGCGCATCGAGCTGGAATGCTGTACGTTGACGTGACATACAGGCTTT	960
Db	901	TC	CACGGTTTCGCGCATCGAGCTGGAATGCTGTACGTTGACGTGACATACAGGCTTT	960
QY	961	GCG	GCTGATGACCGAAGAGAACTGAAGCTGCGCCGTAAGCGCGCGCTTAACGTGTGTAA	1020
Db	961	GCG	GCTGATGACCGAAGAGAACTGAAGCTGCGCCGTAAGCGCGCGCTTAACGTGTGTAA	1020
QY	1021	AA	AGTGATGATCAACACGTGTCTTTGACATCTCGACGCGCGGCACTGCTCTTAATCG	1080
Db	1021	AA	AGTGATGATCAACACGTGTCTTTGACATCTCGACGCGCGGCACTGCTCTTAATCG	1080
QY	1081	GCA	AAATGCCCCGCAAGCTGGGTGACCGCTTAATGTTGCGTCAACAGGAGTACCTCAC	1140
Db	1081	GCA	AAATGCCCCGCAAGCTGGGTGACCGCTTAATGTTGCGTCAACAGGAGTACCTCAC	1140
QY	1141	AA	ACGGCTGAAGGGGATTCGCGCCCGGTAAACCACTCGAACAGCTGATGATTTGTG	1200
Db	1141	AA	ACGGCTGAAGGGGATTCGCGCCCGGTAAACCACTCGAACAGCTGATGATTTGTG	1200
QY	1201	GCG	CACCTGGAAGCGGTGCACTGGGTAGTGTGTTGAAGAGAACGCGCGCAGCGCTTG	1260
Db	1201	GCG	CACCTGGAAGCGGTGCACTGGGTAGTGTGTTGAAGAGAACGCGCGCAGCGCTTG	1260
QY	1261	AT	CGCCGGATCTTCCAGATCTGCTGTGAAGCGCGCGACTATTAACCAAGAAAGATT	1320
Db	1261	AT	CGCCGGATCTTCCAGATCTGCTGTGTGAAGCGCGCGACTATTAACCAAGAAAGATT	1320
QY	1321	GCG	GGAATGAAGTCTGGCGCAACGCTGGCGGAAGTTGTGTCTCACTTTGAAGAC	1380

Accession	Sequence	Position
D6	GCCTGGAGTAAAGATCTCTGGCCACGGTGGCGAATGTTGGGCTCCACTTTGAAGAC	1380
Oy	GGTTGCTCGACGACCAATCATCATGAGAGATCCAAAGGATTAATAAAGGCTTA	1434
D6	GGTTGCTCGACGACCAATCATCATGAGAGATCCAAAGGATTAATAAAGGCTTA	1434

## RESULT 2

```

US-09-912-020-165
Sequence 165, Application US/09912020
Patent No. US20020045592M1
GENERAL INFORMATION:
APPLICANT: Zvekind, Judith
APPLICANT: Ohlsen, Karl L.
APPLICANT: Trawick, John
APPLICANT: Foreyeh, R. Allyn
APPLICANT: Froelich, Jamie M.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
FILE REFERENCE: ELITRA.001DVI
CURRENT APPLICATION NUMBER: US/09/912,020
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: 09/492,709
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 60/117,405
PRIOR FILING DATE: 1999-01-27
NUMBER OF SEQ ID NOS: 485
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 165
LENGTH: 1434
TYPE: DNA
ORGANISM: E. Coli
US-09-912-020-165

```

Query Match	100.0%	Score 1434;	DB 3;	Length 1434;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1434;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	TTGAAAGTAAACGCTGCCAGAGTTTGAACGTCGACGAGATGATGAGTGTGGGATGATG	60
Db	1	ATGAAAGTAAACGCTGCCAGAGTTTGAACGTCGACGAGATGATGAGTGTGGGATGATG	60
QY	61	CTGATCGTTACTGTGTACGCGCCCAACGATGTATCTCGCCGAAAGCCCGCTG	120
Db	61	CTGATCGTTACTGTGTACGCGCCCAACGATGTATCTCGCCGAAAGCCCGCTG	120
QY	121	GTTAAGTGAATACCATCGAAGAACGTCGCGGCGCGCGCTAACGTCGCGATGAATATC	180
Db	121	GTTAAGTGAATACCATCGAAGAACGTCGCGGCGCGCGCTAACGTCGCGATGAATATC	180
QY	181	GCTTCTCTCGGTGCTAATGACGCGCTGATCGGTTGACGCGCATTTGACGATGACGCGC	240
Db	181	GCTTCTCTCGGTGCTAATGACGCGCTGATCGGTTGACGCGCATTTGACGATGACGCGC	240
QY	241	GCGCTGAGTAATCTCTGGCCGACGTCGACGTCMAATGCACTTCGTTCTGTACCGACG	300
Db	241	GCGCTGAGTAATCTCTGGCCGACGTCGACGTCMAATGCACTTCGTTCTGTACCGACG	300
QY	301	CATCGACCATTTACCAATTACGGGTACTTTCCGCGAACCAACAGCTGATTCGACTGGAT	360
Db	301	CATCGACCATTTACCAATTACGGGTACTTTCCGCGAACCAACAGCTGATTCGACTGGAT	360
QY	361	TTTGAAGAAAGTTTCGAAGGTGTGATCCGACGCGTCGACGAGCGGATTAATCAGCGC	420
Db	361	TTTGAAGAAAGTTTCGAAGGTGTGATCCGACGCGTCGACGAGCGGATTAATCAGCGC	420
QY	421	CTGAGTTTGATTTGGCGCGCTGTGCTTTCTGACATACGCCMAAGTTCGCTGGCAAGCTTA	480
Db	421	CTGAGTTTGATTTGGCGCGCTGTGCTTTCTGACATACGCCMAAGTTCGCTGGCAAGCTTA	480



QY 481 CAGCAGATGATCAACTGCGCGGTAAGCGGGGTTTCCGGGCTGATGATCCAAAGT 540  
 DB 481 CAGCAGATGATCAACTGCGCGGTAAGCGGGGTTTCCGGGCTGATGATCCAAAGT 540  
 QY 541 ACCGATTTTGAAGCGCTACCGCGGCTACCGCTGTTAAACCGGAACTCTCGAATTGAA 600  
 DB 541 ACCGATTTTGAAGCGCTACCGCGGCTACCGCTGTTAAACCGGAACTCTCGAATTGAA 600  
 QY 601 GCTGTTGCTGATTAATGATGACCGAAGAGAGATTGTTGAGCGCGGCTGAACTGAT 660  
 DB 601 GCTGTTGCTGATTAATGATGACCGAAGAGAGATTGTTGAGCGCGGCTGAACTGAT 660  
 QY 661 GCGGATTAAGCAACTCTGGGCTCTGTTAGTGAACCGGTTCCGACAGGGGTAATGCTGCTG 720  
 DB 661 GCGGATTAAGCAACTCTGGGCTCTGTTAGTGAACCGGTTCCGACAGGGGTAATGCTGCTG 720  
 QY 721 CAAACCGGATTAAGCGCGCTGCTGATGATGCAACCGAAGCGGAGAGATGATGATGATGAT 780  
 DB 721 CAAACCGGATTAAGCGCGCTGCTGATGATGCAACCGAAGCGGAGAGATGATGATGATGAT 780  
 QY 781 GATGCGGCGGACACGCTGATGATGCGGCTCTGCGGCAACGCTGCGAGCGGATTAATGCTG 840  
 DB 781 GATGCGGCGGACACGCTGATGATGCGGCTCTGCGGCAACGCTGCGAGCGGATTAATGCTG 840  
 QY 841 GAAAGAGCTGCTCTTCTTCCAAATGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 DB 841 GAAAGAGCTGCTCTTCTTCCAAATGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 901 TCCACGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
 DB 901 TCCACGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
 QY 961 GCGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1020  
 DB 961 GCGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1020  
 QY 1021 AAGGATGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1080  
 DB 1021 AAGGATGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1080  
 QY 1081 GCAAAATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140  
 DB 1081 GCAAAATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140  
 QY 1141 AAGGATGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1200  
 DB 1141 AAGGATGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1200  
 QY 1201 GCGGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1260  
 DB 1201 GCGGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1260  
 QY 1261 ATGCGCGGATCTTGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 DB 1261 ATGCGCGGATCTTGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 QY 1321 GCGGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1380  
 DB 1321 GCGGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1380  
 QY 1381 GCTGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1434  
 DB 1381 GCTGCTGATGACCGAAGAGAACTGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1434

RESULT 3  
 US-09-815-242-6204  
 Sequence 6204, Application US/09815242  
 Patent No. US20020061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haelebeck, Robert  
 APPLICANT: Ohlsen, Karl L.

APPLICANT: Zykand, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard  
 TITLE OF INVENTION: Identification of Essential Genes in  
 TITLE OF INVENTION: Prokaryotes  
 FILE REFERENCE: EUTRA.011A  
 CURRENT APPLICATION NUMBER: US/09/815,242  
 PRIOR FILING DATE: 2001-03-21  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 NUMBER OF SEQ ID NOS: 1410  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 6204  
 LENGTH: 1434  
 TYPE: DNA  
 ORGANISM: Escherichia coli  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)...(1434)  
 US-09-815-242-6204

Query Match 100.0%; Score 1434; DB 3; Length 1434;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGTAACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 60  
 DB 1 ATGAAAGTAACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 60  
 QY 61 CTGATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 DB 61 CTGATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 121 GTTAAAGTAACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 180  
 DB 121 GTTAAAGTAACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 180  
 QY 181 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 DB 181 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 241 GCGCTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 DB 241 GCGCTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 301 GCGCTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 DB 301 GCGCTGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 QY 361 TTGAAAGAGTTCGAAGGTTGATTCGCAACCGCTGCAACGAGGCTGATTCGCTGAT 420  
 DB 361 TTGAAAGAGTTCGAAGGTTGATTCGCAACCGCTGCAACGAGGCTGATTCGCTGAT 420  
 QY 421 CTGAGTTGATTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 DB 421 CTGAGTTGATTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 QY 481 CAGCAGATGATCAACTGCGCGGTAAGCGGGGTTTCCGGGCTGATGATCCAAAGT 540

```

Db      481 CAGGATGATGATCAACGCGCGGTAAGCGGGTGTCCGGGTGATGATCCAAAAGT 540
Qy      541 ACCGATTTTGAAGCGCTACCGGGGGCTACGCTGTAAAGCCGATCTCTCGGAATTTGAA 600
Db      541 ACCGATTTTGAAGCGCTACCGGGGGCTACGCTGTAAAGCCGATCTCTCGGAATTTGAA 600
Qy      601 GCTGTGTGCGGTAAATGATGAAGAGAGAGATTTGTGAGCGCGCATGAACTGATT 660
Db      601 GCTGTGTGCGGTAAATGATGAAGAGAGATTTGTGAGCGCGCATGAACTGATT 660
Qy      661 GCCGATTACGAATCTCGGCTCTGTAGTGAACCCGTTCCGAACGSGTATGTCGCTG 720
Db      661 GCCGATTACGAATCTCGGCTCTGTAGTGAACCCGTTCCGAACGSGTATGTCGCTG 720
Qy      721 CAACCGGGTAAAGCGCGCTGCGATATGCGAACCCGAACGCGAGAGATGATGATTAAC 780
Db      721 CAACCGGGTAAAGCGCGCTGCGATATGCGAACCCGAACGCGAGAGATGATGATTAAC 780
Qy      781 GGTGCGGGCGACGAGTATGAGCGTCTGCGCGGCAACGCTGCGAGCGGGTATTCGCTG 840
Db      781 GGTGCGGGCGACGAGTATGAGCGTCTGCGCGGCAACGCTGCGAGCGGGTATTCGCTG 840
Qy      841 GAAGAGCGCTGCTTTCTTTGCCAATGCGCGCGCTGCGCTGCTGCTGCTGCTGCTG 900
Db      841 GAAGAGCGCTGCTTTCTTTGCCAATGCGCGCGCTGCGCTGCTGCTGCTGCTGCTG 900
Qy      901 TCCGACGTTTCCGCGCTGCGATGCGATGCGAATGCTGTAGCTGGAAGTGTGAGATCA 960
Db      901 TCCGACGTTTCCGCGCTGCGATGCGATGCGAATGCTGTAGCTGGAAGTGTGAGATCA 960
Qy      961 GCGCTGATGACCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db      961 GCGCTGATGACCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy      1021 AAAGTGTGATGACCAACGCTGCTTTGACATCTCTGACAGCGCGGCAAGCTCTTAT 1080
Db      1021 AAAGTGTGATGACCAACGCTGCTTTGACATCTCTGACAGCGCGGCAAGCTCTTAT 1080
Qy      1081 GCAAAATGCGCGCGAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db      1081 GCAAAATGCGCGCGAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Qy      1141 AAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db      1141 AAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Qy      1201 GCGCGACCTGGAAGCGCTGCGAGTGTGCTTTGAAGAGAGAGAGAGAGAGAGAG 1260
Db      1201 GCGCGACCTGGAAGCGCTGCGAGTGTGCTTTGAAGAGAGAGAGAGAGAGAGAG 1260
Qy      1261 ATCCCGCGGAGCTGCGAGATCTGCTGTGAAAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db      1261 ATCCCGCGGAGCTGCGAGATCTGCTGTGAAAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy      1321 GCGCGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db      1321 GCGCGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy      1381 GGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1434
Db      1381 GGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1434

```

RESULT 4  
 US-10-282-122A-6430  
 ; Sequence 6430, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Habelbeck, Robert

```

; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EPIIIRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/151,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6430
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-10-282-122A-6430

Query Match      100.0%; Score 1434; DB 7; Length 1434;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAGTAAAGCTGCGCAAGATTGAACGTCAGAGATGATGATGATGATGATGATG 60
Db      1 ATGAAGTAAAGCTGCGCAAGATTGAACGTCAGAGATGATGATGATGATGATGATG 60
Qy      61 CTGATCGTTACTGTAACGCGCCCAACAGTCTATCTCGCGGAAGCGCGGCGCGTG 120
Db      61 CTGATCGTTACTGTAACGCGCCCAACAGTCTATCTCGCGGAAGCGCGGCGCGTG 120
Qy      121 GTTAAAGTAAATCCATCAAGAAAGTCCGGGCGCGCGCTTAAAGTGGAGATGAATATC 180
Db      121 GTTAAAGTAAATCCATCAAGAAAGTCCGGGCGCGCGCTTAAAGTGGAGATGAATATC 180
Qy      181 GCTTCTCTGAGTCTATGACGCTGATCGAGTTGACGGGCAATTGACATGACAGCGCG 240
Db      181 GCTTCTCTGAGTCTATGACGCTGATCGAGTTGACGGGCAATTGACATGACAGCGCG 240
Qy      241 GCGGTGATTAATCTGCGCGACGTCACAGTCAAAATGCGACTTCTGTTCTGTAACGAG 300
Db      241 GCGGTGATTAATCTGCGCGACGTCACAGTCAAAATGCGACTTCTGTTCTGTAACGAG 300
Qy      301 CATCCGACCTTACCAATTAAGGGATCTTCCCGCAACCAAGAGTGAATCCGCTGGAT 360
Db      301 CATCCGACCTTACCAATTAAGGGATCTTCCCGCAACCAAGAGTGAATCCGCTGGAT 360
Qy      361 TTGAAGAGAGTTTGAAGGTTGATCGGACGCGCTGACAGAGCGAGTTAATCAGGCG 420
Db      361 TTGAAGAGAGTTTGAAGGTTGATCGGACGCGCTGACAGAGCGAGTTAATCAGGCG 420

```

```

Qy 421 CTGAGTTGATGCGCGCTGCTTCTGACCTACCGCAAGGTCGCGCAAGGTA 480
Db 421 CTGAGTTGATGCGCGCTGCTTCTGACCTACCGCAAGGTCGCGCAAGGTA 480
Qy 481 CAGAGATGATCAACTGCGCGCTGCTTCTGACCTACCGCAAGGTCGCGCAAGGTA 540
Db 481 CAGAGATGATCAACTGCGCGCTGCTTCTGACCTACCGCAAGGTCGCGCAAGGTA 540
Qy 541 ACCGATTTGAGCGCTACCGCGCGCTACCGCTGTTAAACCGCAATCTCTCGAAATTTGAA 600
Db 541 ACCGATTTGAGCGCTACCGCGCGCTACCGCTGTTAAACCGCAATCTCTCGAAATTTGAA 600
Qy 601 GCTGTGTCGCTAAATGTAAGACCGAAGAAAGATGTTGAGCGCGCAATCTCTCGAAATTTGAA 660
Db 601 GCTGTGTCGCTAAATGTAAGACCGAAGAAAGATGTTGAGCGCGCAATCTCTCGAAATTTGAA 660
Qy 661 GCCGATTAACGAACTCTGCGCTCTGTTAGTGAACCGCTTCCGAAACGAGGATGTCGCTG 720
Db 661 GCCGATTAACGAACTCTGCGCTCTGTTAGTGAACCGCTTCCGAAACGAGGATGTCGCTG 720
Qy 721 CAACCGGCTAAAGCGCGCTGCTGATATGCAACCGCAAGCGCGAAGTGTATGACGTTACC 780
Db 721 CAACCGGCTAAAGCGCGCTGCTGATATGCAACCGCAAGCGCGAAGTGTATGACGTTACC 780
Qy 781 GGTGCGGCGCAACAGGTGATGCGGCTGCTGCGCGCAACCGCTGCGAGCGGCTAATTCGCTG 840
Db 781 GGTGCGGCGCAACAGGTGATGCGGCTGCTGCGCGCAACCGCTGCGAGCGGCTAATTCGCTG 840
Qy 841 GAAAGACCTGCTCTTCTTTCGCAATGCGCGCTGCGCTGCTGCGCAACCGCGCAAC 900
Db 841 GAAAGACCTGCTCTTCTTTCGCAATGCGCGCTGCGCTGCTGCGCAACCGCGCAAC 900
Qy 901 TCACCGGCTTTCGCGCTGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 TCACCGGCTTTCGCGCTGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 GCGCTGATGACCGAAGGAACTGAAGTGGCGCTGAGCGCGCGCAACCGCGTAAACGCGTGA 1020
Db 961 GCGCTGATGACCGAAGGAACTGAAGTGGCGCTGAGCGCGCGCAACCGCGTAAACGCGTGA 1020
Qy 1021 AAGGTGATGATCAACGCTGCTTCTGCAATCTGCAACCGCGCGCAACGCTCTTATCTG 1080
Db 1021 AAGGTGATGATCAACGCTGCTTCTGCAATCTGCAACCGCGCGCAACGCTCTTATCTG 1080
Qy 1081 GCAAAATGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 GCAAAATGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1141 AAGGTGATGATCAACGCTGCTTCTGCAATCTGCAACCGCGCGCAACGCTCTTATCTG 1200
Db 1141 AAGGTGATGATCAACGCTGCTTCTGCAATCTGCAACCGCGCGCAACGCTCTTATCTG 1200
Qy 1201 GCGCACTGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 1201 GCGCACTGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Qy 1261 ATGCGCGGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 ATGCGCGGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Qy 1321 GCGCGGATGATGAAAGTGTGCGCGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Db 1321 GCGCGGATGATGAAAGTGTGCGCGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Qy 1381 GGTGCTGCAACGCAATCATCATCAAGAAATCCCAACGATTAAGAAAGCTAA 1434
Db 1381 GGTGCTGCAACGCAATCATCATCAAGAAATCCCAACGATTAAGAAAGCTAA 1434

```

RESULT 5  
 US-10-771-241-165  
 ; Sequence 165, Application US/10771241  
 ; Publication No. US20040241715A1

```

; GENERAL INFORMATION:
; APPLICANT: Zyklich, Judith
; APPLICANT: Foryth, R. Allyn
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ELITRA.001C
; CURRENT APPLICATION NUMBER: US/10/771,241
; PRIOR APPLICATION NUMBER: 09/492,709
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; NUMBER OF SEQ. ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 165
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: E. Coli
US-10-771-241-165

Query Match 100.0%; Score 1434; DB 8; Length 1434;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAGTAAAGCTGCGCAAGTGTGAACGTCGAGAGTGAATGCTGCTGATGATGATG 60
Db 1 ATGAAAGTAAAGCTGCGCAAGTGTGAACGTCGAGAGTGAATGCTGCTGATGATGATG 60
Qy 61 CTGAGATGTTACTGATGACCGCGCCACCAAGTCTGATCTGCGCGCAACCGCGCGCTG 120
Db 61 CTGAGATGTTACTGATGACCGCGCCACCAAGTCTGATCTGCGCGCAACCGCGCGCTG 120
Qy 121 GTTAAAGTAAATCAATCGAAGAACTCCGCGCGCGCGCTTAACTGCGGATGAAATATC 180
Db 121 GTTAAAGTAAATCAATCGAAGAACTCCGCGCGCGCGCTTAACTGCGGATGAAATATC 180
Qy 181 GCTTCTCTGCTGCTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 GCTTCTCTGCTGCTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy 241 GCGCTGATGATTAATCTGCGCGCAACGTCGCAATGCGCAATGCGCTTCTGTAACGACG 300
Db 241 GCGCTGATGATTAATCTGCGCGCAACGTCGCAATGCGCAATGCGCTTCTGTAACGACG 300
Qy 301 CATCCGACATTAACCAATTAAGGCTTCTCCGCAACCAACGCTGATCCGCTGAT 360
Db 301 CATCCGACATTAACCAATTAAGGCTTCTCCGCAACCAACGCTGATCCGCTGAT 360
Qy 361 TTTGAAGAGGTTTCAAGGCTGATCCGCAACGCTGCAACGCGGATTAATCAAGGCG 420
Db 361 TTTGAAGAGGTTTCAAGGCTGATCCGCAACGCTGCAACGCGGATTAATCAAGGCG 420
Qy 421 CTGAGTTCGATGAGCGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 CTGAGTTCGATGAGCGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 CAGCAATGATCAACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 CAGCAATGATCAACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 ACCGATTTGAGCGCTACCGCGCGCTACCGCTGTTAAACCGCAATCTCTCGAAATTTGAA 600
Db 541 ACCGATTTGAGCGCTACCGCGCGCTACCGCTGTTAAACCGCAATCTCTCGAAATTTGAA 600
Qy 601 GCTGTGTCGCTAAATGTAAGACCGAAGAAAGATGTTGAGCGCGCAATCTCTCGAAATTTGAA 660
Db 601 GCTGTGTCGCTAAATGTAAGACCGAAGAAAGATGTTGAGCGCGCAATCTCTCGAAATTTGAA 660
Qy 661 GCCGATTAACGAACTCTGCGCTCTGTTAGTGAACCGCTTCCGAAACGAGGATGTCGCTG 720
Db 661 GCCGATTAACGAACTCTGCGCTCTGTTAGTGAACCGCTTCCGAAACGAGGATGTCGCTG 720
Qy 721 CAACCGGCTAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

```

```

Db      721  CAAACGGGTAAAGCGCGCTCATATGCAACCAAGCGCAGAAAGTATGACGTAC 780
Qy      781  GGTGCGGGCGCAGCGGTATTTGGGCTCTGCGCGCAACGCTGGCAGGGGAAATTCCTG 840
Db      781  GGTGCGGGCGCAGCGGTATTTGGGCTCTGCGCGCAACGCTGGCAGGGGAAATTCCTG 840
Qy      841  GAAGAAGCTGCTCTTCTTGGCAATGCGCGCGCTGGCGGTGTGTCTGGCAAACTGGAA 900
Db      841  GAAGAAGCTGCTCTTCTTGGCAATGCGCGCGCTGGCGGTGTGTGTCTGGCAAACTGGAA 900
Qy      901  TCACGCGTTTGGCGGATGACCTGGAAAATCTGTACGTGAGCTGCAATACAGGCTTT 960
Db      901  TCACGCGTTTGGCGGATGACCTGGAAAATCTGTACGTGAGCTGCAATACAGGCTTT 960
Qy      961  GCGCGTATGACCGGAAGGAACCTGAAGTGGCGGTAGCGGAGCGCGTAAACGTGGTAA 1020
Db      961  GCGCGTATGACCGGAAGGAACCTGAAGTGGCGGTAGCGGAGCGCGTAAACGTGGTAA 1020
Qy      1021  AAAGTGTATGACCAACGCTGTCTTTGACATCTCTGACGCGCGGCACTCTCTTATCTG 1080
Db      1021  AAAGTGTATGACCAACGCTGTCTTTGACATCTCTGACGCGCGGCACTCTCTTATCTG 1080
Qy      1081  GCAAAATGCGCGCAAGCTGGGTGACCGCTTGTATGTTGCGCTCAACAGCGATGCTCA 1140
Db      1081  GCAAAATGCGCGCAAGCTGGGTGACCGCTTGTATGTTGCGCTCAACAGCGATGCTCA 1140
Qy      1141  AAAAGGCTGAAAGGGGAAATCCCGCGCGGTAAACCACTCGCAACGCGGTAAATTTG 1200
Db      1141  AAAAGGCTGAAAGGGGAAATCCCGCGCGGTAAACCACTCGCAACGCGGTAAATTTG 1200
Qy      1201  GCGCGCATGGAAGCGGTGACCTGGGTAGTGTCTTTGGAAGAGACAGCGCGCAGCGCTG 1260
Db      1201  GCGCGCATGGAAGCGGTGACCTGGGTAGTGTCTTTGGAAGAGACAGCGCGCAGCGCTG 1260
Qy      1261  ATCGCGCGGATCTTGGCAGATCTGTGTGTAAGAGCGCGCGCACTTAAACGAGAGATT 1320
Db      1261  ATCGCGCGGATCTTGGCAGATCTGTGTGTAAGAGCGCGCGCACTTAAACGAGAGATT 1320
Qy      1321  GCCCGGAGTAAAGAGTCTGGGCGCAAGGTGGCGAGATGTTGGCTCACTTTGAAGAC 1380
Db      1321  GCCCGGAGTAAAGAGTCTGGGCGCAAGGTGGCGAGATGTTGGCTCACTTTGAAGAC 1380
Qy      1381  GGTGTCTGACGACCAATCATCATGAAGATCCACAGATTAAGGCTAA 1434
Db      1381  GGTGTCTGACGACCAATCATCATGAAGATCCACAGATTAAGGCTAA 1434

```

RESULT 6  
US-10-282-122A-20018  
Sequence 20018, Application US/10282122A  
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyckind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: EITRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

```

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 7814
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 20018
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-10-282-122A-20018

```

Query Match 71.0%; Score 1018.2; DB 7; Length 1428;  
Best Local Similarity 82.2%; Pred. No. 0;  
Matches 1170; Conservative 0; Mismatches 253; Indels 0; Gaps 0;

```

Qy      1  ATGAAAGTAAACGCTGCCAGAGTTTGAACGTGACGAGAGTATGTTGATGATG 60
Db      1  ATGAAAGTAAACGCTGCCAGAGTTTGAACGTGACGAGAGTATGTTGATGATG 60
Qy      61  CTGATCGTACTAGTACGCGCCACCAAGTCGTATCTGCGCGGAAGCGCGTCCGCTG 120
Db      61  CTGATCGTACTAGTACGCGCCACCAAGTCGTATCTGCGCGGAAGCGCGTCCGCTG 120
Qy      121  GTTAAAGTAAATACATCGAAGACGTCCGCGCGCGCGCTTAACTGGCGATATATC 180
Db      121  GTTAAAGTAAATACATCGAAGACGTCCGCGCGCGCGCTTAACTGGCGATATATC 180
Qy      121  GTTAAAGTAAATACATCGAAGACGTCCGCGCGCGCGCTTAACTGGCGATATATC 180
Db      121  GTTAAAGTAAATACATCGAAGACGTCCGCGCGCGCGCTTAACTGGCGATATATC 180
Qy      181  GCTTCTCTGCGTCTAATGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db      181  GCTTCTCTGCGTCTAATGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Qy      181  GCTTCTCTGCGTCTAATGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db      181  GCTTCTCTGCGTCTAATGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Qy      241  GCGCTGAGTAATCTGCGCGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 300
Db      241  GCGCTGAGTAATCTGCGCGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 300
Qy      301  CATCCGACATTAACCAATTTACGCGGTACTTCCCGCAACCAACGCTGATCCGTGGAT 360
Db      301  CATCCGACATTAACCAATTTACGCGGTACTTCCCGCAACCAACGCTGATCCGTGGAT 360
Qy      361  TTTGAAGAGGTTTCAAGGTGTTGATCCGACGCGCTGACGAGGAGTTATCAAGCGG 420
Db      361  TTTGAAGAGGTTTCAAGGTGTTGATCCGACGCGCTGACGAGGAGTTATCAAGCGG 420
Qy      421  CTGAGTTCAATTGCGCGCGTGGTCTTCTGATACGCGCAAGGTGCGTGGCAAGGCTA 480
Db      421  CTGAGTTCAATTGCGCGCGTGGTCTTCTGATACGCGCAAGGTGCGTGGCAAGGCTA 480
Qy      481  CAGCAGATATTCMACTGGCGGTAAAGCGGTGTTCCGCTGCTGATTCGCAAAAGT 540
Db      481  CAGCAGATATTCMACTGGCGGTAAAGCGGTGTTCCGCTGCTGATTCGCAAAAGT 540
Qy      541  ACCGATTTTGAAGCGCTACCGCGCGCTAGCTGTAAACCGCAATCTCTCGGAATTTGA 600
Db      541  ACCGATTTTGAAGCGCTACCGCGCGCTAGCTGTAAACCGCAATCTCTCGGAATTTGA 600
Qy      601  GCTGTTGCGGTAAATGTAAGACCGAAGAGATGTTGAAGCGCGCATGAAGTGAATT 660
Db      601  GCTGTTGCGGTAAATGTAAGACCGAAGAGATGTTGAAGCGCGCATGAAGTGAATT 660

```

661 GCCGATTACGAACTCTGGCTCTGTAGTACCGGTTCCGACAGGATATGCTGCTG 720  
661 GCGGATTACGAACTCTGGCTCTGTAGTACCGGTTCCGACAGGATATGCTGCTG 720  
721 CAACCGGATTAAGCGCGCTGATATGCAACCCAGCGGAGAGATGATGACGTTACC 780  
721 CAGCGCGGAGAAAGCGCGCTGATATGCAACCCAGCGGAGAGATGATGACGTTACC 780  
781 GGTGCGGCGGACACGCGGATTTGGGCTCTGCGCGGCAACGCTGAGCGGATTAATGCTG 840  
781 GGTGCGGCGGATTAAGCGGATTTGGGCTCTGCGCGGCAACGCTGAGCGGATTAATGCTG 840  
841 GAAGAACTCTGCTCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 900  
841 GAAGAACTCTGCTCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 900  
901 TCACCGGCTTCTGCGGATCTGAGCTGGAACGCGGCTGCGGCTGCGGATTAATGCTG 960  
901 TCACCGGCTTCTGCGGATCTGAGCTGGAACGCGGCTGCGGCTGCGGATTAATGCTG 960  
961 GCGGCTGATGACCGGAGGAACTGAACTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
961 GCGGCTGATGACCGGAGGAACTGAACTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
1021 AAAGTGTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTG 1080  
1021 AAAGTGTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTG 1080  
1081 GCAATGCGCGGCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1140  
1081 GCAATGCGCGGCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1140  
1141 AAAGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTG 1200  
1141 AAAGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTG 1200  
1201 GCGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1260  
1201 GCGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1260  
1261 ATGCGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTG 1320  
1261 ATGCGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTG 1320  
1321 GCGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1380  
1321 GCGGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1380  
1381 GGTGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1423  
1381 GGTGCTGATGACCAACGCTGCTTCTTGGCAATGCGCGGCTGCTGCTGCTGCTGCTGCTG 1423

## RESULT 7

US-09-815-242-9661  
Sequence 9661, Application US/09815242  
Patent No. US2002061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zykend, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: EUTRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
PRIOR APPLICATION NUMBER: 2001-03-21  
PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9661  
LENGTH: 1434  
TYPE: DNA  
ORGANISM: *Salmonella typhi*  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1434)  
US-09-815-242-9661  
Query Match 70.3%; Score 1008.6; DB 3; Length 1434;  
Best Local Similarity 81.8%; Pred. No. 0;  
Matches 1164; Conservative 0; Mismatches 259; Indels 0; Gaps 0;  
1 ATGAAAGTAAACGCTGCGGAGATTGAACTGCAAGAGTATGATGATGATGATGATGATG 60  
1 ATGAAAGTAAACGCTGCGGAGATTGAACTGCAAGAGTATGATGATGATGATGATGATG 60  
61 CTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120  
61 CTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120  
121 GTTAAAGTAAACGCTGCGGAGATTGAACTGCAAGAGTATGATGATGATGATGATGATG 180  
121 GTTAAAGTAAACGCTGCGGAGATTGAACTGCAAGAGTATGATGATGATGATGATGATG 180  
121 GTTAAAGTAAACGCTGCGGAGATTGAACTGCAAGAGTATGATGATGATGATGATGATG 180  
181 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
181 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
181 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
241 GCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
241 GCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
301 CATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
301 CATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
301 CATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
361 TTTGAAAGGCTTGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
361 TTTGAAAGGCTTGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
421 CTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
421 CTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
421 CTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
481 CAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
481 CAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
481 CAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
541 ACCGATTTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
541 ACCGATTTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
601 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
601 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
661 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
661 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720

Db 661 GCCAATTACGACCTTTCCGCGCTGTGTGTCAGCGGCTTCCGAAAGGGAATGACGCTGCTG 720  
 Qy 721 CAACCGGGTAAAGCGCGCTGCATATGCAACCAAGCGCAGAAGTGTATGACGTTACC 780  
 Db 721 CAACCGGAATTAAGCGCGCTGCATATGCGCAACGAGCGCAGAAGTGTATGACGTTACC 780  
 Qy 781 GGTGCGGGCGCAGCAGCTGATTTGGCGCTCTGGCGCGCAACGCTTGGCAAGCGGGTAAATTCGCTG 840  
 Db 781 GGTGCGGGCGCAGTATGCTGATCGCGTGTGCGCGCGCAGCAGCTTGGCGCGGGAATTAACCTTG 840  
 Qy 841 GAAGAAGCTCTCTCTTTCCAAATGCGCGCGCTGCGCGGTGCGGTGCGGAAACCTGGAAAC 900  
 Db 841 GAGAGAGCGGTGTATTTTCGCAATGCGCGCGCGCGCGCGGTGCGGTGCGGAAACCTGGAAAC 900  
 Qy 901 TCCACGCTTTGCGCGCATGCACTGCAAAATCTGTATGCTGACGCTGCAAGATACAGGCTTT 960  
 Db 901 TCAACGCTTTCCCTATTTAGCTGGAATAACGATGCGCGCAACGCGCGGATACCGGCTTC 960  
 Qy 961 GCGGTGATGACCGAAGAAAGTGAAGCTGCGCGGTAGCGGCGAGCGCGTAAACGTGTGAA 1020  
 Db 961 GCGGTGATGACCGAAGAAAGTGAAGCTGCGCGGTAGCGGCGAGCGCGTAAACGTGTGAA 1020  
 Qy 1021 AAAGTGTGATGACCAACAGGTGTCTTTGACATCTCTGACGCGCGGCAACGCTCTTATCTG 1080  
 Db 1021 AAAGTGTGATGACCAACAGGTGTCTTTGACATCTCTGACGCGGCGCAACGCTCTTATCTG 1080  
 Qy 1081 GCAATATGCGCGCAAGCTGCTGATGCGCTTGAATTTGCGCGTCAACAGCGATGCTTCAAC 1140  
 Db 1081 GCGAATGCGCGCAAGCTGCTGATGCGCTTGAATTTGCGCGTCAACAGCGATGCTTCAAC 1140  
 Qy 1141 AAACGCGCTGAAAGGAGATTTCCGCGCGGTAAACCACTGCAACGCGTATGATTTGCTG 1200  
 Db 1141 AAACGCTGAAAGGAGATTTCCGCGCGGTAAACCACTGCAACGCGTATGATTTGCTG 1200  
 Qy 1201 GCGCGCATGGAAGCGCTGATGCTGCTTTGAAAGGAGATGACGCGCGCAACGACTG 1260  
 Db 1201 GCGCGCATGGAAGCGCTGATGCTGCTTTGAAAGGAGATGACGCGCGCAACGACTG 1260  
 Qy 1261 ATCCGCGGATCTTGCAGATCTGCTGTGTAAGAGCGCGCGCATTAACCGAAGATTT 1320  
 Db 1261 ATTGCGGATCTTGCAGATCTGCTGTGTAAGAGCGCGCGCATTAACCGAAGATTT 1320  
 Qy 1321 GCGCGAGTAAAGAGTCTGCGCGCAACGCTGCGCAAGTGTGTCTCAACTTTGAAAC 1380  
 Db 1321 GCGCGAGTAAAGAGTCTGCGCGCAACGCTGCGCAAGTGTGTCTCAACTTTGAAAC 1380  
 Qy 1381 GCTTGTGATGACCAACATCATCAAGAAAGTCCAAACAGATA 1423  
 Db 1381 GCTTGTGATGACCAACATCATCAAGAAAGTCCAAACAGATA 1423

## RESULT 8

US-10-282-122A-39294  
 ; Sequence 39294, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl  
 ; APPLICANT: Zyklich, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: EUTPA 034A  
 ; CURRENT APPLICATION NUMBER: US/10/282.122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 39294  
 ; LENGTH: 1434  
 ; TYPE: DNA  
 ; ORGANISM: *Salmonella typhi*  
 ; US-10-282-122A-39294

Query Match 70.3%; Score 1008.6; DB 7; Length 1434;  
 Best Local Similarity 81.8%; Pred. No. 0;  
 Matches 1164; Conservative 0; Mismatches 259; Indels 0; Gaps 0;  
 Qy 1 ATGAAGTAAACGCTGCGCAGAGTTTGAACGTGACAGAGTATGCTGTGTAATGATG 60  
 Db 1 ATGAAGTAAATATGCAAGGTTTGAACGTGACAGAGTATGCTGTGTAATGATG 60  
 Qy 61 CTGATGCTTACTGTGTAAGCGCCCAACAGTCTATCTGCGCGAAGCGCGCTG 120  
 Db 61 CTGATGCTTACTGTGTAAGCGCCCAACAGTCTATCTGCGCGAAGCGCGCTG 120  
 Qy 121 GTTAAGTAAATCATCAAGAAAGTCCGCGCGCGCGCTTAAACGCGAATATC 180  
 Db 121 GTTAAGTAAATCATCAAGAAAGTCCGCGCGCGCGCTTAAACGCGAATATC 180  
 Qy 181 GCTTCTCTGAGTCTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Db 181 GCTTCTCTGAGTCTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Qy 241 GCGCTGATTAATCTGTGCGCGAGCTCAACGTCAATGCAATGCACTTCTGTGACGAG 300  
 Db 241 GCGCTGATTAATCTGTGCGCGAGCTCAACGTCAATGCAATGCACTTCTGTGACGAG 300  
 Qy 301 CATCCGACCTTTCGAAATTAACGCGTATCTTCCGCAACCAACGCGTATCCGCTGAT 360  
 Db 301 CATCCGACCTTTCGAAATTAACGCGTATCTTCCGCAACCAACGCGTATCCGCTGAT 360  
 Qy 361 TTTGAAGAAAGTTCGAAAGTGTGATTCGCAAGCGCTGCAAGCGGATTAATCAGGCG 420  
 Db 361 TTTGAAGAAAGTTCGAAAGTGTGATTCGCAAGCGCTGCAAGCGGATTAATCAGGCG 420  
 Qy 421 CTGAGTTCGATTTGCGCGCTGCTGCTTCTGATGACGCAAGGTCGCTGCGAAGCTA 480  
 Db 421 CTGAGTTCGATTTGCGCGCTGCTGCTTCTGATGACGCAAGGTCGCTGCGAAGCTA 480  
 Qy 481 CAGCAGATGATCAACTGCGCGCTGTAAGGCGGCTTCCGCTGCTGATGATCCAAAGGT 540  
 Db 481 CAGCAGATGATCAACTGCGCGCTGTAAGGCGGCTTCCGCTGCTGATGATCCAAAGGT 540  
 Qy 541 ACCGATTTGAGCGCTACCGCGCGCTACGCTGTAAACCGCAATCTCTGGAATTTGAA 600  
 Db 541 ACCGATTTGAGCGCTACCGCGCGCTACGCTGTAAACCGCAATCTCTGGAATTTGAA 600  
 Qy 601 GCTGTGTGCTGTAATGTAAGACGGAAGAGATTTGAGCGCGCATGAATGAT 660



Db	601	GGGGTGGGGGAAATGTAAAGCGAAGCAACTGTTGAACGGCGATGAACTCATTT	660
Qy	661	GGCGATTACGAACCTCGGCTCTGTTAGTACCCGTTCCGAACAGGGTAATGTCCGTCTG	720
Db	661	GGCGATTACGAACCTTTCCGCGCTGTGTGTCACGCGTTCGAAACAGGGAATGACGCTGCTG	720
Qy	721	CAACCGGGTAAAGGCGCGGTGATATGCAACCAGGCGGAGGAATGTATGACGTTACC	780
Db	721	CAACCGAATTAAGCCCGCTACATATGCGAGCGAGCGGAGGAATTTATGATGTAAAC	780
Qy	781	GGTGGGGGCGACACCGGTATTTGCGCTCTGCGCGCAACGCTGCGACGCGGTAAATTCGCTG	840
Db	781	GGTGGGGGCGATACCGGTATTCGCGCTGTCTGCGCGGAGCGCTGTGCGCGGAATATCCCTG	840
Qy	841	GAAGAACTGTCTTCTTTTCCCATGCGGCGGCTGGCGGTGTGTGCGCAAACTGGAACC	900
Db	841	GAGGAGGCGGTGTATTTTCCCATATCGGCGGCGGCGGTATGATGTAACTCGGAGCG	900
Qy	901	TCCACGGTTTGGCCGATCGAGCTGTGAATAATGTGTACGTGTGACGTGTGAGATACAGGCTTT	960
Db	901	TCCACGGTTTCCCTTATTTAGGTGTGAATAACGATGTGCGCGACGCGCGGAATACCGGCTTC	960
Qy	961	GAGCGTATGACCGAAGAGAACTGAAGCTGTGGCGGTAGCGGCGGACGCGGTAAAGTGTGTAA	1020
Db	961	GGCGTATGACCGAAGAGAACTTTGAGACAGGCCGCTGTGCGACGCGCGGTAAAGGTGTGTAA	1020
Qy	1021	AAAGTGTGTATGACCAACGAGTGTCTTTGACATCTGTGCAACGCGGCGACGCTCTTATCTG	1080
Db	1021	AAAGTGTGTATGACCAACGAGGCTTTTGCATATTTGTGCAACGCGGCGCACGCTCTTATCTG	1080
Qy	1081	GCAATATCCCGCAAGCTGGGTGTACCGCTTATTTGTTGCCGTTCACAGCGATGCTCCACC	1140
Db	1081	GCGAATCGGCGCGAATCTGGGCGACCGCGCTGATTTGTGTGCGTCAATATGTAGACCGCTCGACT	1140
Qy	1141	AAACGGCTGAAAAGGGGAATTTCCCGCGGTAAACCACTCGAACAGCGATGATTTGTGCTG	1200
Db	1141	AAACGTCTGAAAAGGGGAATGCGGTTCGTTATTCGCTCGAACAGCGATGATGCTGCTG	1200
Qy	1201	GAGCGATCTGAAGCGGTGTGACTGGGTAGTGTGTGTGAAGAGAACCGCGCAGCGCTTG	1260
Db	1201	GAGCGGCTGTGAAGTGTGTGACTGGGTGTGTCTTTTGAAGAGAAATCCGCCACAGACTG	1260
Qy	1261	ATCGCGGGAATCTTCCAGATCTGCTGTGAAGGCGCGCATATTAACAGAAAGATT	1320
Db	1261	ATTCGCGGTATTTCTGCGCGATCTGCTGTAAAGGCGCGCATATTAACCGGAAGAGATC	1320
Qy	1321	GCCGGGAATTAAGAGTCTGGGCGCAACGCTGTGCGAAAGTGTGGTGTCTCAACTTTGAAGAC	1380
Db	1321	GCGGCGAACGAAAGAGTCTGGGCGCAACGCGCGCAAGTCAATGTGTCTAACTTGAAAGAT	1380
Qy	1381	GTTTGCTCGACGACCAATCATCAAGAAATCCAAAGAGATA	1423
Db	1381	GTTTGTTCACGACCAATCATCAAAAGATCCAGACCGAGA	1423
RESULT 9			
US-10-282-122A-24093			
Sequence 24093, Application US/10282122A			
Publication No. US20040029129A1			
GENERAL INFORMATION:			
APPLICANT: Wang, Liangsu			
APPLICANT: Zamudio, Carlos			
APPLICANT: Malone, Cheryl			
APPLICANT: Hagelbeck, Robert			
APPLICANT: Ohlsen, Karl			
APPLICANT: Zyekind, Judith			
APPLICANT: Wall, Daniel			
APPLICANT: Trawick, John			
APPLICANT: Carr, Grant			
APPLICANT: Yamamoto, Robert			
APPLICANT: Foreyth, R.			
APPLICANT: Xu, H.			

```

1 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
2
3 FILE REFERENCE: EITRA 034A
4
5 CURRENT APPLICATION NUMBER: US/10/282,122A
6
7 PRIOR FILING DATE: 2003-02-20
8
9 PRIOR APPLICATION NUMBER: 60/191,078
10
11 PRIOR FILING DATE: 2000-03-21
12
13 PRIOR APPLICATION NUMBER: 60/206,848
14
15 PRIOR FILING DATE: 2000-05-23
16
17 PRIOR APPLICATION NUMBER: 60/237,727
18
19 PRIOR FILING DATE: 2000-05-26
20
21 PRIOR APPLICATION NUMBER: 60/230,335
22
23 PRIOR FILING DATE: 2000-09-06
24
25 PRIOR APPLICATION NUMBER: 60/230,347
26
27 PRIOR FILING DATE: 2000-09-09
28
29 PRIOR APPLICATION NUMBER: 60/242,578
30
31 PRIOR FILING DATE: 2000-10-23
32
33 PRIOR APPLICATION NUMBER: 60/253,625
34
35 PRIOR FILING DATE: 2000-11-27
36
37 PRIOR APPLICATION NUMBER: 60/257,931
38
39 PRIOR FILING DATE: 2000-12-22
40
41 PRIOR APPLICATION NUMBER: 60/267,636
42
43 PRIOR FILING DATE: 2001-02-09
44
45 PRIOR APPLICATION NUMBER: 60/269,308
46
47 PRIOR FILING DATE: 2001-02-16
48
49 Remaining SEQ Application data removed - See File Wrapper or PALM.
50
51 NUMBER OF SEQ ID NOS: 76614
52
53 SOFTWARE: PatentIn version 3.1
54
55 SEQ ID NO 24093
56
57 LENGTH: 1431
58
59 TYPE: DNA
60
61 ORGANISM: Klebsiella pneumoniae
62
63 US-10-282-122A-24093

```

Query Match	70.2%	Score 1007	DB 7	Length 1431
Best Local Similarity	81.7%	Pred. No. 0		
Matches 1163	Conservative	0	Mismatches 260	Indels 0
			Gaps	0
QY	1 ATGAAGTAAAGCGCTGECAGAGTTTGAAGGTGACAGAGTGAATGTGGTGTGATGTGAATG	60		
DB	1 ATGAAGTAAAGCGCTGECAGAGTTTGAAGGTGACAGAGTGTGGTGTGATGTGAATG	60		
QY	61 CTGGATCGTTACTGGTACGCGCCCAACAGTGTATCTGCGCGGAAGCGCGGTGCCGTG	120		
DB	61 CTGAGCCCTCTCTGGTACGCGCCCAACAGTGTATTTCCCGGAAGCCCGGTGCCGTG	120		
QY	121 GTTAAAGTAAATCCATGATGAAGAAAGTCCGCGCGCGCGCGCTTAAAGTGGCGATGAATATC	180		
DB	121 GTAAAGGTGAATAATATGAAAGAAAGTCTGCGCGCGCGCGCAACGTACGATGAACATTT	180		
QY	181 GCTTCTCTCGGTGTAAATGACAGCGCTGGTTCGGGTGTGACGGGCAATTGACGATGACGGCCG	240		
DB	181 GCTCTGCTGGGGGCAACTCGCGCTTGGTGGGATTTGACCGGATTTGATACCGCGCCCG	240		
QY	241 GCGCTGAATAATCTCTGCGCGAGCTCAACGTCAATGCGACTTGGTTTCTGTACCGACG	300		
DB	241 GCGCTGAGCCAGCGCTGCGCAACGTCAATGGAATGCGACTTGGTCTCCGTCGCAACT	300		
QY	301 CATCGACCAATTACCAATTAACGGTACTTTCCCGCAACCAACAGTGAATCCGTCTGAAT	360		
DB	301 CACCGCATTAATCACCAACGTGGGGTGTCTGCGCAATCAACAGTGAATCCGCTCGAC	360		
QY	361 TTGGAAGAAAGTTTGAAGTGTATCCGACAGCGCTGACAGACGGAATTATACAGGCG	420		
DB	361 TTGGAAGAGGCTTCTCGCGGTGATTCGACAGCGGATGACAGAGGCAATTCAGCAGGGG	420		
QY	421 CTGAGTTGATTTGCGCGCTGTGCTTTCTGACTACGCCAANAGTGCCTGGCAACGGTA	480		
DB	421 CTGGGCTTATTTGGCGCGCTGTGCTTTCTGACTATGCCAANAGGCGCTGACAGCGTC	480		
QY	481 CAGCAGATGATCCAACTGGCGCGTAAACCGGGTGTCCGGTGTGATTTGAATCCAAAGGT	540		
DB	481 CAGACATGATCCGACTGCAACGAGAGGGCGTGGCGGTGTGATTCATTCGAAAGGC	540		



QY	541	ACGCAATTTGAGCGCTACCGGGGGGGCTACGGCTGTTTAAAGCGCAATCTCTCGAAATTTTGA	600
Db	541	ACCGATTTTCGAAACGCTATCCGGGGCGACAGTTCCTCAACCCGAACTCTCTGAATTTGA	600
QY	601	GCTGTGTGCGTAAATGTTAAGACCGAGAAAGAGATTGTGAGCGCGCATGAATCTGATTT	660
Db	601	GCGGTGGGGCAAGTGCACAGATGAAGCGCAGATCTGTTGAGCCGCGCATMACTCAATC	660
QY	661	GGCGATTACGAACCTCGGGCTCTGTGTATGACCCGTTCCGAACAGGGGATATGCGTGGTG	720
Db	661	GCCGAAATTCGAACCTGCGGCGCTGTGTGACCGGCTTGTAGCAGGGGATACCGTGTGCTG	720
QY	721	CAACCGGTTAAAGCGCGCTGCATATGCGCAACCCAAAGCGCAGAAAGTGTATGACGTTAC	780
Db	721	CAGCGGGGGCGTCCGGCGCTGCATATATGCGAACCCAGCGCCAGAAAGTGTATGACGTGACC	780
QY	781	GGTGCGGGCGCACACGCTGTATTTGGCGTCTCTGGCGCGCAACGCTGCGAGCGGGTAAATTCG	840
Db	781	GGCGCCGGCGCACACGCTGTATTTGGCGTGTGTGGCGCGAACCTTGGCGGTCCGGCAACACCT	840
QY	841	GAAAGAGCGCTACTTCTTTCGCAATGCGCGGCTGGCGTGGTGGTTCGCGCAAACTGGAAAC	900
Db	841	GAAAGAGCGCTCTCTATTTTCGCTTAACCGCGCGCGCGCGCTGTGTGTTCGTAACTGGGTACA	900
QY	901	TCCAACGTTTCGCGCATCGACCTGAAATGCTGTACGTGACCTGCAGATACAGGCTTT	960
Db	901	TCCAACGTTTCCTCTGTGTGACCTGAAACCGCGGTCCGCGCGCGGTGCAGAACCGGCTTT	960
QY	961	GGCGGTATGACCGGAAGAGAACTGAAAGCTGACGTCGATGAGGCGGACGCGCTTAAACGTGTGAA	1020
Db	961	GGCGGTATGACCGAGAGAGAAAGCTCAAAACGCGGTGGTGGCGGGCGCGCAACGCGGCGAG	1020
QY	1021	AAAGTGTGTATGACCAACCGATGTCTTTTGCATCTCTGCACGCGCGGCAAGTCTTATCTG	1080
Db	1021	AAAGTGTGTATGACCAACCGGCGTCTTTCGACATTTCTGCATGCGGTCATGTCTCTATCTG	1080
QY	1081	GCAATGCCCCCGCAAGCTGGGTGACCGGCTTGTATTTGTGCGCTCAACGAGATGGCTCCACC	1140
Db	1081	GCGAATGGCGCGCAACGCTGGGTGACCGGTCTGTATCTGGGGGTCAACAGGAGATGCTCCACC	1140
QY	1141	AAAACGCTGAAAGAGGGGATTTCCCGCCCGGTAAACCACTCGAACAGCTATGATTTGTGCTG	1200
Db	1141	AAACGCTCTGAAAGAGGGGAACCGCGCCGGTGAACCCGCTGGAACAGCGGATATATGTGTGCTG	1200
QY	1201	GACGCACTGGAAGCGGTGCACTGGGTATGTGTCTTTGAAGAGACACGCGCGACGCGCTTG	1260
Db	1201	GGGGCGCTGGAAGCGGTGCACTGGGTGTCTCTTTCGAAGAAAGATATCTCCGACGGCGCTG	1260
QY	1261	ATGCGCGGGAATCTTGCGCAGATCTGTGTGTGAAGAGCGCGCACTTAAACCAAGAGATTT	1320
Db	1261	ATGCTGTGGCATCTCTCCCGGATGTGTCTGGTCAAGGGCGCGCACTAACACCGGACAGATTC	1320
QY	1321	GCGCGAGTAAAGAGTCTGGGCGAACGCTGGCGAAGTGTGTGTCTCAACTTTGAAGAC	1380
Db	1321	GCGGCGACGGAAGTCTGGGCGAACGCTGGCGAAGTGTGTGTCTCAACTTTGAAGAT	1380
QY	1381	GGTGTCTGACGACCAATCATCTCAAGAAAGATCCCAACGAGTAA	1423
Db	1381	GGGTCTCAACTCAATTTATTTAAGAAATAACAAAGAGTA	1423

```

RESULT 0
US-10-282-122A-36984
Sequence 36984, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel

```

```

/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.03A4
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 36984
/ LENGTH: 1397
/ TYPE: DNA
/ ORGANISM: Salimella paratyphi A
/ US-10-282-122A-36984

```

Query Match	Best Local Similarity	85.2%	Score 935.2	DB 7	Length 1397
Matches 1137	Conservative	61.1%	Pred. 0.61e-298	Mismatches 255	Indels 7
				Gaps	
QY	22	TTTGAACGTGACGAGTGTATGCTGGTTGGTATGTATGATCTGGATCGTTACTGTACGGC	81		
DB	1	TTTGAACGTGACGAGCGTATGTTGGTGTATGTAAATGCTGAGATCGCTATGTATGGTGC	60		
QY	82	CCCAACAGTGGTATCTCGCGGGAAGCCGCGGTCGCCGCTTAAAGTATACATCGCA	141		
DB	61	CCCACTTGGCGTATTTCAACCGAAGCCGCGTCCCGTGTAAAGTAAATACCTTTAG	120		
QY	142	GAACGTCCGGGCGGCGCGGCTTAACGTGGCGATGAATATCGCTTCTTCGCTGCTAATGCA	201		
DB	121	GAACGTCCGGGCGGCGCGGCGCAACGTGGCGATGAACATTTGGTCTCTCGGAGCGAAGCGC	180		
QY	202	CGCTCGTGGGGTTTGAAGGGGATTTAGCATGACAGCGCGCGCTGATTAATCTCTGGCGC	261		
DB	181	CGTCTGGTGGCGCTGACCGGGTATTTGATGATCGC--CGCGCGGCTGACGAAACGTTGGCG	238		
QY	262	GACGTCAACGTCAATGCGACTTCTGTTCTGTACCGACGCATTCGACCATTAACCAATTA	321		
DB	239	GAGGTCAATGGAAGTGGCACTTGGTTCTGTGCGGACGCATTCGACCATTAACCAACTG	298		
QY	322	CGGGTACTTTCCCGAACCAACAGCTATTCGCTCGGATTTTGAAGAGGTTTGGAGGT	381		
DB	299	CGCGGTACTACAGTATCAAGTACAGTCAATTCGTTGATTTTGAAGAGGCTTTGAGGGG	358		
QY	382	GTTGATCCGACAGCGCTGCACAGCGGATTAATCAGCGCTGAGTTGATTTGGCCGCGCTG	441		
DB	359	GTTGACCGCGACGCGTTTGCATGAGGTATCAACAGAGGCGCTGGGATCGATCGCGCGCTG	418		
QY	442	GTCCTTTTGTACTACGCGCAAGGTCGCTGGCGACGCTTACAGAGATGATTCATATTCGGG	501		
DB	419	GTATTGTCCGATTAAGCCAAAGCGCTCTGACACAGCGTGCAGACTATGATTTTCCCTTAGG	478		

QY 502 CGTAAAGCGGCTGTTCCGTCGCTGATGATCCAAAGGTAACCGATTTTGAAGCCTACCGC 561  
 DB 479 CGCCAGGGGGCGGCGGCTGCTCATGATCCGAAGGAAGGATTTTGAAGCCTTACCGC 538  
 QY 562 GGGCCTACGCTGTTAAGCCGAATCTCTCGGAATTTTGAAGCTGTTGCTGTAATGTAG 621  
 DB 539 GGGCCACGCTGCGAGCGCCAAACCTTCTGAAATTCAGAGCGGCTTGGGGGAATGTAA 598  
 QY 622 ACCGAAGAGATTTGTAGCGCGGCAAGAACTGATTCGATTCGAATCTCTGCT 681  
 DB 599 AGCCAAAGCAATCGTTGAAACGGGCAATGAAGCTCATTCGATTCGATTCGCTG 658  
 QY 682 CTGTAGTGAACCGCTTCCGAACAGGATATGCTGCTGCAACCGGCTTAAAGCGCCCTG 741  
 DB 659 CTGTGCTGACGCGCTTCCGAACAGGATATGCTGCTGCAACCGGCTTAAAGCGCCCTA 718  
 QY 742 CATATGCCAACCCAAAGCGAGAAAGTATGACCTTACCGGTCGGGCGACACAGCTGAT 801  
 DB 719 CATATGCCAACCCAAAGCGAGAAAGTATGACCTTACCGGTCGGGCGATACAGCTGATC 778  
 QY 802 GGGCTCTGGGCGCAACGCTGCGACGCGGATATTCGCTGGAAGAGCCTGCTTCTTCC 861  
 DB 779 GGGCTCTGGGCGCGAC-CTGCGCGCGGAAATACCTGGAAGAGCGCTGTTATTCGCC 837  
 QY 862 AATGCGGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921  
 DB 838 AATGCGGCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897  
 QY 922 CTGGAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 981  
 DB 898 CTGGAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954  
 QY 982 CTGGAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041  
 DB 955 TTGAGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014  
 QY 1042 GTCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101  
 DB 1015 GTCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074  
 QY 1102 GACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161  
 DB 1075 GACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134  
 QY 1162 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221  
 DB 1135 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194  
 QY 1222 TTGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281  
 DB 1195 TTGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1254  
 QY 1282 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1341  
 DB 1255 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1313  
 QY 1342 GCCAAGCGTGGCGAAGTGTGTGCTCACTTTGAAGACGCTTCTGACGACCAATC 1401  
 DB 1314 GCCAAGCGTGGCGAAGTGTGTGCTCACTTTGAAGACGCTTCTGACGACCAATC 1373  
 QY 1402 ATCAAGAGATTCACAGAGAT 1423  
 DB 1374 ATCAAGAGATTCACAGAGAT 1395

RESULT 11  
 US-10-282-122A-41981  
 Sequence 41981, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Karl  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 PRIOR FILING DATE: 2003-02-20  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR FILING DATE: 2001-02-16  
 PRIOR FILING DATE: 2001-02-16  
 REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: Patent version 3.1  
 SEQ ID NO 41981  
 LENGTH: 1431  
 TYPE: DNA  
 ORGANISM: Yersinia pestis  
 US-10-282-122A-41981

Query Match 61.3%; Score 879.2; DB 7; Length 1431;  
 Best Local Similarity 76.2%; Pred. No. 2.1e-279; Indels 0; Gaps 0;  
 Matches 1082; Conservative 0; Mismatches 338;

QY 1 ATGAAGTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60  
 DB 1 ATGAAGTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60  
 QY 61 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 DB 61 TTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 QY 121 GTTAAAGTAAATCCATTCGAAGAAAGTCCGCGCGCGCGCTTAAAGTGAAGTAAATC 180  
 DB 121 GTTAAAGTAAATCCATTCGAAGAAAGTCCGCGCGCGCGCTTAAAGTGAAGTAAATC 180  
 QY 181 GCTTCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 DB 181 GCTTCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 QY 241 GCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 DB 241 GCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 301 CATCCGACATTAACCAATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 DB 301 CATCCGACATTAACCAATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 361 TTGGAAGAGGTTTCAAGAGTGTGATCCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

```

Db      361  TTGAGAGAGGCTTTGACGCTGTGATCCACGCCGATATTGAGCCGATTCACATGCGC 420
Qy      421  CTGAGTTGANTGGCGCTGTGTCTTTGACACGCAAGGTGGCTGGCAAGGTA 480
Db      421  TTGCAACAGATTGGCGCTGTGTCTGTCCATTATGCAAGGGGCAATTGACAGGTA 480
Qy      481  CAGCAGATGATCCAACTGCGCTGTAAGCGGGGTTCGCGTGTGATTTGACAAAGT 540
Db      481  CAGCGCATGATCCAACTGCGCACTGTAAGCGGATGTCCCGTTTGTGATTTGACAAAGC 540
Qy      541  ACCGATTTTGAAGCGCTACCGCGGCTACGCTGTTAAAGCCGAATCTCTGGAATTGAA 600
Db      541  AGTCAATTTGAGCCCTACCGGTGTGCAATTTGTTGACCGCAATTTGTCTGAATTTGAA 600
Qy      601  GCTGTCTCGGTAATGTAAGACGGAAGAGATGTTGTGAGCGCGCATGAACTGAT 660
Db      601  GCAGTGTGAGGCGCTGTAAATGAAAGAGATGTTGTTAACCGCGGTATGCAATGCTGT 660
Qy      661  GCCGATTAACGAATCTCGGCTCTGTGATGACCCGTTCCGAACAGGGTATGCTGCTG 720
Db      661  GCAGATTTTGAATCTTTCAGCATTTGCTGTGACCGGTTCTGTAACAGGCAATGACCTTGT 720
Qy      721  CAACCGGGTAAAGCGCGCTGCTGATATGCAACCAAGCGCAGAAAGTGTATGACGTTAC 780
Db      721  CAACCTGGGTAAACCGCGCTACATTTGCGCATGAGGCAAAAGAGTGTGATGATGACC 780
Qy      781  GGATCGGGCGCAGCGGTGATTGGGCTCTGTGCGCAACGCTGAGAGCGGGTAATTCGCTG 840
Db      781  GGATCTGCGCATGACGATTTATTTGGCTGTGTGCTGCGCGCTGTGAGGTAACAGCTTG 840
Qy      841  GAAGAGACCTGCTCTTTGCAATGCGCGGCTGCGGTGTGATGCGCAACTGAGAAC 900
Db      841  GAAGAGATCATGTTTCTGTGCTTAATGCTGCGCGCGGTGTGATGATGAGTAAATTTGGATAC 900
Qy      901  TCCACGTTTTGCGGATGAGCTGGAATAATCTGTAGTGTGACGTGCAATGACGCTTT 960
Db      901  TCCACGTTTTCTCCGATTTGAGTGTGAAACCGCATTCGCTGCGCGGTGAGAAACCGGTTTT 960
Qy      961  GGCGTGTATGACGGAAGGAACTGAAGCTGCGCGGTGAGCGCGCGGTAAACGTGTGAA 1020
Db      961  GGCGTGTATGATGAGGAACTGAAGTACGCGGTGAGCGCGCGGTGAGCGCGCGGTGAGGAG 1020
Qy      1021  AAGGTGTATGACCAACGATCTTTGACATCTCCATGACGAGCAATGTTCTTATCTG 1080
Db      1021  AAGGTGTATGACCAACGATCTTTGACATCTCCATGACGAGCAATGTTCTTATCTG 1080
Qy      1081  GCAATGCTGCAAGCTGAGTGTCTTTGACATCTCTGACGCGCGGACGCTCTTATCTG 1140
Db      1081  GCAATGCTGCAAGCTGAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1140
Qy      1141  AAGCGGTGAAAGGGGATTTCCCGCGGTAAACCACTGCAACGCGATGATGTTGTGCTG 1200
Db      1141  AAGCGGTGAAAGGGGATTTCCCGCGGTAAACCACTGCAACGCGATGATGTTGTGCTG 1200
Qy      1201  GGCGCATGGAAGCGCTGACCTGTGTGTTGAAGAGACAGCGCGCAGCGCTTG 1260
Db      1201  GGCGCATGGAAGCGCTGAGTGTGTGTTGAAGAGATGAGCGCGCAGCGCTTG 1260
Qy      1261  ATCCCGCGGATCTTGCGCATCTGTGTGTAAGCGCGCGCATTTAAACGAAAGATTT 1320
Db      1261  ATTCCTGATTTATTTGGCGCATCTGTGTGTAAGCGCGCGCATTTAAAGCGCGCAATTT 1320
Qy      1321  GCCCGAGTAAAGAGTCTGGGCGCAACGCTGCGCAAGTGTGTGCTCACTTTGAAGAC 1380
Db      1321  GCAAGTAAAGAGTCTGGGCGCGCGCGGTGAGGTTAAAGGTTAGTCAATTTTGAAGAT 1380
Qy      1381  GGTTGCTGACGACCAATCATCAAGAGATCCAAAG 1420
Db      1381  GGTTGCTGACCAATCATCAATTAATTAATCAATCAAAAG 1420

```

RESULT 12  
US-10-282-122A-32536

```

; Sequence 32536, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32536
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-10-282-122A-32536

Query Match 50.6%; Score 725; DB 7; Length 1422;
Best Local Similarity 69.4%; Pred. No. 2.1e-228;
Matches 986; Conservative 0; Mismatches 435; Indels 0; Gaps 0;

Qy      1  ATGAAGTAAAGCTTGCAGAGTTTGAACGTGACGAGTATGTTGTGTGTGATGATG 60
Db      1  ATGAAGTAAAGCTTGCAGAGTTTGAAGCAATGTGTGTGTGTGTGTGTGTGTGTGT 60
Qy      61  CTGATCTGTTACTGTATGAGCGCCCAACAGTCTATCTGCGCGGAAGCGCCGCTG 120
Db      61  TTGAGCGTTTATGTGTATGTCTTACAGTGTGATCTTCAACAGACCTGTGTGCTG 120
Qy      121  GTTAAAGTAAATCAATGAAAGAGTCCGGGCGCGCGGTAAAGTGTGCGATGATATC 180
Db      121  GTTAAAGTAAATCAATGAAAGAGTCCGGGCGCGCGGTAAAGTGTGCGATGATATC 180
Qy      181  GCTTCTCTGCTGTATGACAGCCCTGTGTGTGACGCGCAATTTGACAGCGCGC 240
Db      181  GCTTCTTGTAGGGCTTAATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Qy      241  GCGCTGAGTAAATCTGTGCGCGAGCTCAAGCTCAAAATGCACTTGTGTGTGTGTGT 300
Db      241  GCATTAGCAATCAATTAAGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      301  CATCCGACATTAACAAATTAAGGTTATTTCCGCAACCAAGCTGATCCGTGTGAT 360

```

[illegible]

257 TGGCCGACGTCAAGTCATGCAATGCACTTCTGTTCTGTATCCGACGACATCCGACATTAACA 316  
257 TAGAAAAACAAAAATGATTTGTTGATTTAGAAACCAATCAACATTTTCTA 316  
317 AATTACGGGTACTTCCCGCAACACAGCTGATCCGCTGCGATTTTGAAGAAGTTTCG 376  
317 AATTACGATTTTATCTGTCATCAACAGCTGCTCCGCTTGATTTGAAAGAAATTTCA 376  
377 AAGGTTGATCCGACGCGCTGCAACGAGGAAATTAATCAGCGCTGATTCGATTTGGCG 436  
377 ATAAATGATGATTCAGAGATTTATTAAGCAATTAAGAAAGTCGCTGAAATTAACGTTG 436  
437 CGCTGCTCTTCTGACTACCCAAAGGTGCGCTGCGACGCGTACAGAGATGATCCAC 496  
437 CTTTGAATCTTTCTGATTAAGGCAAGGCAAGCTTAAGATTTCAAGAAATGATTTCAA 496  
497 TGGCGCGTAAAGCGGGTGTCCGCTGCTGATTTCCAAAGATTCGATTTTGAAGGCT 556  
497 TTGCAACGCAAGCAATGCTGCTGTTGATCGATCCAAAGGAACTGATTTTGAAGGCT 556  
557 ACCGCGGCGCTAACGCTGTTAAAGCGCAATCTCTGCGAAATTTGAAGCTGTTGGTAAAT 616  
557 ATCGTGGGCTACATTAATGACCAACCAATATGCTGAATTTGAAGCTGTTGGGTAAT 616  
617 GTAAGACGGAAGAGATTTGTGACGCGCATGAATGATTTGCCGATTCGACACTCT 676  
617 GCAATACGGAAGAGATTTGTAAGAAAGTTTAAATTAATTTCTGATTTGAATTA 676  
677 CGGCTCTGTTAGTGAACCCGTTCCGAACAGGGTATGTCGCTGCTGCAACCGGCTAAAGCC 736  
677 CCGCATTTTGTGATCGCGTCTGTAAGAAAGCAATGATTTACGCCCAATCAAGAC 736  
737 CGCTGCAATATGCAACCAACGCGAGAGATGATGATTAACGCTGCGGCGGCAACGCG 796  
737 CTTATCATTTTCCCACTGTTGCAAAAGATGTTGATGATGACGAGACCGGCTGACATG 796  
797 TGATTTGCGCTCTGCGCGCAACGCTGCGAGCGGATTTCTGCGAAGAACCTGCTTCT 856  
797 TCATTAAGGATTTAGCAACCGCATTAAGAGATGACGTTCTTTGAAAGAACTTGTATACC 856  
857 TTGCGCATGCGCGCGCTGCGTGTGTGCGGCAAACTGCGAACTTCCACGTTTCCGCGA 916  
857 TAGCCAAATGTTGCGCGCAAGATTTGTGTGCGGAAATTTGCGCACTTCAACGCTTTCACCG 916  
917 TCGAGCTGGAATATGCTGTACGTGACGTGACAGATTAACGCTTTGCGCTGATGACGAG 976  
917 TCGAATCTTGAATATGCGATTAATGCTGCTCTGTAACCTGGAATTTGCGCATTTATGATGAG 976  
977 AGGAATGAGCTGCGCTGACGCGGACGCGTAAACGTGTGAAGAAAGTGTGATGACCA 1036  
977 CAGATTTAAAGATGCTGTGCAACAGTAAAGCGCGGTGAAGAAATTTGTATGATGACTA 1036  
1037 ACGGTGCTTTGACATCTGCAACGCGCGGACGCTCTTATCTGCGCAATCCCGCAAGC 1096  
1037 ATGCTGTTTGAATTTTGTGATCCAGGCAATTTCTTATTTGAAATATCAGCGCAAT 1096  
1097 TGGGTGACGCGCTGATTTGTCGCGTCAACAGCGATGCTTCAACCAACGCTGGAAGGCG 1156  
1097 TGGGCGATGCTGATTTGTCGCGTAAACAGCGAGATTTCTGTTAAACGCTTAAAGGTTG 1156  
1157 AAGTGGCCCAATTAATCTTGAAGACGATAGCGGATTTGCTGCTTGTGCGATCCG 1216  
1157 ATTCGCGCCGCGTAAACCACTCGAACAGCGTATGTTGCTGCGCGCACTGGAAGCGG 1216  
1217 TCGACTGGGTAGTGTCTTTGAAGAGACAGCGCGCAAGCTTGTATCCCGGATCTTGC 1276  
1217 TAGACTGGTGTGCTTTCACTGAAGATCAACCAAGTTTATTCGCGCAAAATTTCTAC 1276  
1277 CAGATCTGCTGCTGTAAGAGCGCGGCACTTAAACGGAAGATTTGCGGAGTAAAGAG 1336  
1277 CAGATCTTTTGTGTAAGAGCGCGGCACTTAAACGGAAGATTTGCGGAGTAAAGAG 1336  
1337 TCTGGGCGCAACGCTGCGAGATGTTGCTCAACTTTGAAGAGCGTTGCTGACGACCA 1396

Db 1337 TTGGGCGAAACGCTGCGAGATGTTAAAGTCTAAACTTTGAAATGTTGTTCAACAA 1396  
Qy 1397 ACATCATCAGAGATCCAA 1416  
Db 1397 ATGTGATGAAAAAATTA 1416

RESULT 14  
US-10-329-670-1/c  
Sequence 1, Application US/10329670  
Publication No. US20040018503A1  
GENERAL INFORMATION:  
APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragment of Invention: Thereof, and Uses Thereof  
FILE REFERENCE: P186P1  
CURRENT APPLICATION NUMBER: US/10/329,670  
PRIOR FILING DATE: 2002-12-24  
PRIOR APPLICATION NUMBER: US 09/643,990  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: US 08/487,429  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/426,787  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1830121  
TYPE: DNA  
ORGANISM: Haemophilus influenzae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4747)..(4747)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (9921)..(9921)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (10150)..(10150)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (29298)..(29298)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36543)..(36543)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36551)..(36551)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36636)..(36636)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (40808)..(40810)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44416)..(44416)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44905)..(44905)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature

```
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (53569)..(53569)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119524)..(119524)
```

```
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

```
Query Match 44.34; Score 635.2; DB 7; Length 1830121;
Best Local Similarity 65.94; Pred. No. 2.9e-197;
Matches 922; Conservative 0; Mismatches 478; Indels 0; Gaps 0;
Cy 17 CAGAGTTGAACGCGCAGAGTGAAGTGGTGGTGGTGAATGATCGATCGTTACGCT 76
Db 1596554 CAGATTTAAGCAAGCAAAAGTACTGCTATTAGCGCAGCGATGATCGTTATTCGT 1596495
```



QY 77 ACGGCCCCACCACTGATCTCGCCGGAAGCCCGGTCGCCGTGTTAAAGTAATCA 136  
Db 1596494 TCGGCGCAACCGATATTTTCAACAGAAAGCAACGATACAGTGTTTGTGTAAAGAAA 1596435  
QY 137 TCGAAGAACTGTCGGGCGGCGCGGCTTAAGTGGATGATATGCTTCTCGTCTA 196  
Db 1596434 ATGAAGAAACGCGCGGCGGTCGACAAATGCGGATGAAATATGCTTCACTCATATGAC 1596375  
QY 197 ATGACGCGCTGCGGGGTGCGGACATGACATGACGCGCGCGTGTAAATCTC 256  
Db 1596374 CCGTTCAGTTATAGGATTTGATGACAAAGTAACTGCTTGTGCACTTTCCTTTAT 1596315  
QY 257 TGGCCGACGTCAACGTCAATATGCACTTCTTCTGTAACGACGCACTCCACATTA 316  
Db 1596314 TAGAAAAACAAAAATGATTTGTAATTTTGTGATTAAGAAACCATCAACATTTCTA 1596255  
QY 317 AATTACGGGTACTTTCCCGCAACCAAGCTGATCCGCTGCGATTTTGAAGAAGTTTC 376  
Db 1596254 AATTACGATTTTATCTGTCATCAACGCTGCTCCGCTGTGATTTGAAAGAAATTTCA 1596195  
QY 377 AAGGTGTTGATCCGACGCGCTGACGAGCGGATTTAATCAAGCGCGCTGAGTTGAGCG 436  
Db 1596194 ATAAATGATGATTCAGAGGATTTAATAGCAAGTTAAGAAAGTCGCTGAAATTAACGTTG 1596135  
QY 437 CGCTGCTGCTTCTGACTACGCCAAAGGTGCGCTGCGAAGCTACAGACATGATCCAC 496  
Db 1596134 CTTTGATTTCTTCTGATTAAGCGCAAGGCAAGCTTAAAGATTTCAAGAAATGATTTCAA 1596075  
QY 497 TGGCGCGTAAAGCGGCGTGTCCGCTGCTGATTAATCCAAAGTAAACGATTTTGAAGGCT 556  
Db 1596074 TTGCAACGCAAGCAATGTGCTGTGATGATCCAAAGGCAATGATTTTAAAGGTT 1596015  
QY 557 ACGCGCGGCTACGCTGTTAAAGCGCAATCTCTGGAATTTGAAGCTGTTGCGTAAAT 616  
Db 1596014 ATCGTGGGCTACATTAATTAAGCAACCATATGCTGAATTTGAAGCTGTTGCGTAAAT 1595955  
QY 617 GTAAGACCGAAGAAAGATTTGTGACGCGCATGAATCTGATTCGATTAACGACTCT 676  
Db 1595954 GCAATACGAAAGAAAGATTAATTAAGAAAGTTTAAATTAATTTGATTAATTA 1595895  
QY 677 CCGCTCTGTTAGTACCCGTTCCGAACAGGATATGTGCTGCTGCAACCGGTTAAAGCG 736  
Db 1595894 CCGCACTTTGCTGACGCGTCTGAAAGGCAATTAATTAACGCAATCAAGAC 1595835  
QY 737 CGCTGCAATATGCAACCGCAAGCAAGAAAGTATGAGCTTAACGCTGCGGCGACACG 796  
Db 1595834 CTTATCATTTTCCCACTGTTGCAAAAGAAAGTGTGATGATGACGCGGATACACTG 1595775  
QY 797 TGATTTGCGCTCTGCGCGCAACGCTGCGACGCGGTAATTCGCTGAAAGAAAGCTGTTCT 856  
Db 159774 TCATTTAGGCTATTAAGCAACCGCATTAAGACATGACGTTCTTGAAGAAATCTGTTACC 1595715  
QY 857 TTGCTCAATGCGCGCGCTGCGCTGCTGCTGCAAACTGCGAACTTCCAGCGTTTCGCGA 916  
Db 1595714 TAGCCAAATGTTGCGCAAGAAATGTGTGAGGAAATGAGGAACTTCAACGCTTTCGACCG 1595655  
QY 917 TCGAGCTGAAATATGCTGTACGTGACGTGACATATACAGCTTTGCGCTGATGACCAAG 976  
Db 1595654 TGGAACTTGAATATGCAATGCTGTCTGCAACAGCTAAAGCGCGGTGAAGAAATTTGATGACTA 1595535  
QY 977 AGGAATCGAAGCTGCGCGTACGCGACGCGCTAAACGTGTGAAAGAAAGTGTGATGACCA 1036  
Db 1595594 CAGAATTTAAAGATGCTGTGCAACAGCTAAAGCGCGGTGAAGAAATTTGATGACTA 1595535  
QY 1037 ACGGTGCTTTGACATCTGACGCGCGGCAAGTCTTATATCTGCAAAATGCCCCGCAAGC 1096  
Db 159534 ATGGCTGTTTGAATTTTGAATTCAGGCGCAATTTCTTATTTTGAAGAAATGCAAGCAAT 1595475  
QY 1097 TGGGTAACCGCTGATTTGTGCGCTCAACAGAGATGCTTCAACCAAAACGCGTAAAGGCG 1156  
Db 1595474 TGGGCGATGCTGATATTTGTTGCGGTAAACAGCGAGATTTCTGTTAAAGCTTAAAGGTTG 1595415  
QY 1157 ATTCCCGCCCGTAAACCCACTCGAACAGCGTATGATTTGTGCTGCGCGCACTGGAAGCGG 1216

Db 1595414 AAGTCCCAATTAATTAATCTTGAAGACGTAATGCGGATTTGCGCATCCG 1595355  
QY 1217 TCGACTGGGATGATGCTTTGAAAGACACGCGGACGCTGATGCGCGGATCTTGC 1276  
Db 1595354 TAACTGATGATGCTTCTTCACTGAAGATACCAACAGCTTTAAATGCGGAAATTTCTAG 1595295  
QY 1277 CAGATCTGCTGTAAGAGCGCGCACTATTAACAGAAAGATTTGCGGAGTAAAGAG 1336  
Db 1595294 CAGATCTTTATGTCAAAGCGCGGATTAACAAACCGAAGATTTGACAGGCAATTAAGAG 1595235  
QY 1337 TCTGGGCAACGCTGCGCAAGTGTGCTCACTTTGAAGCGGTTGCTGACACCA 1396  
Db 1595234 TTGGGCAACGCTGCGCAAGTGTGCTCACTTTGAAGCGGTTGCTGACACCA 1595175  
QY 1397 ACATCATCAAGATCCAA 1416  
Db 1595174 ATGTGATGAAAAATTA 1595155  
RESULT 15  
US-10-158-865-1/c  
Sequence 1, Application US/10158865  
Publication No. US20040203093A1  
GENERAL INFORMATION:  
APPLICANT: Pleischmann et al.  
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fragm  
TITLE OF INVENTION: Thereof, and Uses Thereof  
FILE REFERENCE: PB166P2C1D1  
CURRENT APPLICATION NUMBER: US/10/158, 865  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US 09/557,884  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US 08/476,102  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/426,787  
PRIOR FILING DATE: 1995-04-21  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1830121  
TYPE: DNA  
ORGANISM: Haemophilus Influenzae  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (4747)..(4747)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (9921)..(9921)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (10150)..(10150)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (29298)..(29298)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (36543)..(36543)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (36551)..(36551)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (36636)..(36636)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc feature



```

LOCATION: (40808) .. (40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416) .. (44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905) .. (44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975) .. (44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593) .. (45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732) .. (45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036) .. (47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334) .. (51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602) .. (51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786) .. (51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805) .. (51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (53369) .. (53369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309) .. (65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313) .. (65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024) .. (80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091) .. (100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696) .. (102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121) .. (105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107246) .. (107246)
OTHER INFORMATION: n equals a,t,c, or g

OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136) .. (117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750) .. (119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924) .. (119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038) .. (120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344) .. (121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167) .. (122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336) .. (122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340) .. (131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360) .. (131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910) .. (139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398) .. (140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750) .. (142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058) .. (145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171) .. (145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942) .. (145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197) .. (147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841) .. (150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500) .. (152500)
OTHER INFORMATION: n equals a,t,c, or g

```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 18:51:39 ; Search time 938 Seconds  
(without alignments)  
10188.880 Million cell updates/sec

Title: US-10-520-820-28

Perfect score: 1434  
Sequence: 1 atgaagatgaacgtcgcgcaga.....aacagatataaagcctaa 1434

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 4996997 seqs, 333246308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_21:\*

- 1: geneseqn19808:\*
- 2: geneseqn19808:\*
- 3: geneseqn20008:\*
- 4: geneseqn2001a:\*
- 5: geneseqn2001b:\*
- 6: geneseqn2002a:\*
- 7: geneseqn2002b:\*
- 8: geneseqn2003a:\*
- 9: geneseqn2003b:\*
- 10: geneseqn2003c:\*
- 11: geneseqn2003d:\*
- 12: geneseqn2004a:\*
- 13: geneseqn2004b:\*
- 14: geneseqn2005a:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1434	100.0	1434	3	AA65973
2	1434	100.0	1434	4	AA652567
3	1434	100.0	1434	5	AAH81477
4	1434	100.0	1434	8	ACA18560
5	1434	100.0	1434	12	ADH88994
6	1018.2	71.0	1428	8	ACA32148
7	1011.8	70.6	1500	11	ACH96242
8	1008.6	70.3	1434	4	AA656024
9	1008.6	70.3	1434	8	ACA51424
10	1007	70.2	1431	8	ACA36223
11	935.2	65.2	1397	8	ACA49114
12	879.2	61.3	1431	8	ACA54111
13	870.6	60.7	1323	11	ACH96312
14	784.8	54.7	110000	10	ACF65388_02
15	784.8	54.7	110000	10	ACF65388_02
16	783.4	54.6	1425	10	ACF11372
17	725	50.5	1422	8	ACA44666
18	724.2	50.5	1476	10	ADP02177
19	635.2	44.3	1431	8	ACA34479

C	20	635.2	44.3	110000	2	AA42063_15	Continuation (16 o
	21	625.6	43.6	349980	13	ADP05377	Ad05377 Haemophil
	22	617.4	43.1	1431	8	ACA33070	ACA33070 Prokaryot
	23	507.8	35.4	707	6	ABQ21090	Abq21090 Oligonuc
C	24	507.8	35.4	707	6	ABQ21091	Abq21091 Oligonuc
	25	497.8	34.7	1425	4	AA554344	AA554344 Pseudom
	26	497.8	34.7	1425	4	ACA42715	ACA42715 Prokaryot
	27	494.2	34.5	1581	11	ABD17809	ABD17809 Pseudom
	28	493	34.4	1422	8	ACA45258	ACA45258 Prokaryot
	29	465.6	32.5	1407	8	ACA33795	ACA33795 Prokaryot
C	30	425.2	29.7	707	6	ABQ21092	Abq21092 Oligonuc
	31	425.2	29.7	707	6	ABQ21093	Abq21093 Oligonuc
	32	406.2	28.3	2055	5	AA589869	AA589869 DNA encod
C	33	397	27.7	555	6	ABQ45445	Abq45445 Oligonuc
	34	397	27.7	555	6	ABQ45444	Abq45444 Oligonuc
	35	391	27.3	1251	11	ABD17700	ABD17700 Pseudom
C	36	390.4	27.2	1001	4	AA691438	AA691438 Haemophil
	37	390.4	27.2	1001	6	ABK37818	ABK37818 DNA seque
C	38	388.6	27.1	3641	13	ADP05422	Ad05422 Haemophil
	39	350.2	24.4	963	11	ABD17574	ABD17574 Pseudom
C	40	340.8	23.8	555	6	ABQ45443	Abq45443 Oligonuc
	41	340.8	23.8	555	6	ABQ45442	Abq45442 Oligonuc
C	42	302.6	21.1	706	6	ABQ20706	Abq20706 Oligonuc
	43	302.6	21.1	706	6	ABQ20707	Abq20707 Oligonuc
C	44	301	21.0	700	10	ACF66848	ACF66848 Photorhab
	45	286	19.9	948	8	ACA23565	ACA23565 Prokaryot

## ALIGNMENTS

RESULT 1	AAA65973	standard; DNA; 1434 BP.
ID	AAA65973	
XX	XX	
AC	AAA65973;	
XX	XX	
DT	05-OCT-2000	(first entry)
XX	XX	
DE	E. coli proliferation associated coding sequence SEQ ID NO:165.	
XX	XX	
KM	Escherichia coli; E. coli; proliferation; inhibition; screening;	
KW	antimicrobial; bacterial growth; antisense therapy; antibacterial; de.	
XX	XX	
OS	Escherichia coli.	
XX	XX	
PN	NC0200044906-A2.	
XX	XX	
PD	03-AUG-2000.	
XX	XX	
PF	27-JAN-2000; 2000NC-US002200.	
XX	XX	
PR	27-JAN-1999; 99US-0117405P.	
XX	XX	
PA	(BLIT-) ELITRA PHARM INC.	
XX	XX	
PI	Zyckind J, Ohlsen KL, Trawick J, Forsyth RA, Froelich JM;	
PI	Carr GJ, Yamamoto RT, Xu HH;	
DR	WPI; 2000-514822/46.	
XX	XX	
DR	P-PSDB; AAB15966.	
XX	XX	
PT	Novel polynucleotides and polypeptides associated with microorganism	
PT	proliferation, used to identify inhibitors of bacterial growth and	
XX	XX	
PS	Claim 8; Page 146-147; 316pp; English.	
CC	AAA65809 to AAA65889 and AAA66058 to AAA66138 represent nucleotide	
CC	sequences derived from Escherichia coli which inhibit E. coli	
CC	proliferation. AAA65890 to AAA66055 and AAB15886 to AAB16040 represent	
CC	nucleotide and protein sequences associated with E. coli proliferation.	
CC	AAA66056 and AAA66057 represent primers used for sequencing E. coli	

CC proliferation inhibiting nucleotide inserts in an example from the  
CC present invention. Methods from the present invention can be used to  
CC identify a proliferation- required gene in a microorganism, by contacting  
CC a microorganism with a proliferation- required gene activity inhibitory  
CC nucleic acid identified in another organism, and determining if  
CC inhibition occurs in the second microorganism. The nucleic acid sequences  
CC identified as being required for bacterial growth and proliferation, can  
CC be used for antisense therapy for killing bacteria

**Sequence** 1434 BP; 328 A; 352 C; 447 G; 307 T; 0 U; 0 Other;

Query Match	100.0%	Score 1434	DB 3	Length 1434
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1434	0	Mismatches	0	Gaps 0

Qy	1	ATGAAGTAAACGCTGCGACAAGTTTGAACGTGAGGAGAGTGAATGAGTCTTGGTATGATGTAATG	60
Db	1	ATGAAGTAAACGCTGCGACAAGTTTGAACGTGAGGAGAGTGAATGAGTCTTGGTATGATGTAATG	60
Qy	61	CTGGATCGTTACTGATGACGCGCCCAACAGTCGTATCTGCGCGGAAGCGCGGATCCCGTG	120
Db	61	CTGGATCGTTACTGATGACGCGCCCAACAGTCGTATCTGCGCGGAAGCGCGGATCCCGTG	120
Qy	121	GTTAAAGTGAATACATCCAGAAAGTCGCGGCGCGCGCTTAACGTGGCGATGAATATC	180
Db	121	GTTAAAGTGAATACATCCAGAAAGTCGCGGCGCGCGCTTAACGTGGCGATGAATATC	180
Qy	181	GCTTCTCTGGGCTTAATGACGCGCTGGTCGGGTTGAAGGCGATTGAAGCAAGCAAGCGCGC	240
Db	181	GCTTCTCTGGGCTTAATGACGCGCTGGTCGGGTTGAAGGCGATTGAAGCAAGCAAGCGCGC	240
Qy	241	GCGCTGAGTAAATCTCTGCGCGACGTCAAGTCAATGCGACTTCGTCTGTATCCGACG	300
Db	241	GCGCTGAGTAAATCTCTGCGCGACGTCAAGTCAATGCGACTTCGTCTGTATCCGACG	300
Qy	301	CATCCGACCATTAACCAATTACGGGTACTTTCCGCAACCAACGCTGATCCGCTGAT	360
Db	301	CATCCGACCATTAACCAATTACGGGTACTTTCCGCAACCAACGCTGATCCGCTGAT	360
Qy	361	TTTGAAGAAAGTTTGGAGGTGTGAATCCGCGACGCGCTGACGAGCGGATTAACAGAGG	420
Db	361	TTTGAAGAAAGTTTGGAGGTGTGAATCCGCGACGCGCTGACGAGCGGATTAACAGAGG	420
Qy	421	CTGAGTTCGATTGGCGCGCTGCTGCTTCTGACTACGCAAGGTGCGCTGGCAAGCGTA	480
Db	421	CTGAGTTCGATTGGCGCGCTGCTGCTTCTGACTACGCAAGGTGCGCTGGCAAGCGTA	480
Qy	481	CAGCAGATGATCCAACTGCGCGGTAAAGCGGGTGTCCGGTGTGATTTGATCCAAAGGT	540
Db	481	CAGCAGATGATCCAACTGCGCGGTAAAGCGGGTGTCCGGTGTGATTTGATCCAAAGGT	540
Qy	541	ACGGAATTTGAGGCTAACGCGCGCGCTAACGCTTTAACGCGGAACTCTCGGAATTTGAA	600
Db	541	ACGGAATTTGAGGCTAACGCGCGCGCTAACGCTTTAACGCGGAACTCTCGGAATTTGAA	600
Qy	601	GCTGTTCTCGGTAAATGTAAGAACGGAAGAAGATTGTGTAAGCGCGGATGAATGAT	660
Db	601	GCTGTTCTCGGTAAATGTAAGAACGGAAGAAGATTGTGTAAGCGCGGATGAATGAT	660
Qy	661	GCCGATTACGAATCTCTGCGCTCTGTTAGTGAACCGCTTCCGAACAGGATATGTCCCTGTG	720
Db	661	GCCGATTACGAATCTCTGCGCTCTGTTAGTGAACCGCTTCCGAACAGGATATGTCCCTGTG	720
Qy	721	CAACCGGGTAAAGCGCGCGCTGAATAGCCAAACCAAGGCGAGGAAGTATGAAGTTAAC	780
Db	721	CAACCGGGTAAAGCGCGCGCTGAATAGCCAAACCAAGGCGAGGAAGTATGAAGTTAAC	780
Qy	781	GGTGGCGGCGACACGCTGATTTGGCGTCTGGGCGGAACGCTGGCAGCGGGATTAATTCGCTG	840
Db	781	GGTGGCGGCGACACGCTGATTTGGCGTCTGGGCGGAACGCTGGCAGCGGGATTAATTCGCTG	840
Qy	841	GAAGAAGCTGCTTTTGGCAAATGCGGCGCTGGCGTGGTGGCAACTGGGAAC	900
Db	841	GAAGAAGCTGCTTTTGGCAAATGCGGCGCTGGCGTGGTGGCAACTGGGAAC	900

Db	841	GAAGAAACCTGCTTCTTTGGCAATGCGGCGGGCTGGCGTGGTGGTTCGGCAATCTGGAAAC	900
Qy	901	TCACGGTTTCCGCCGATCGAGCTGGAAAAATGCTGTACGTGGACGTGCAGATCAAGGCTTT	960
Db	901	TCACGGTTTCCGCCGATCGAGCTGGAAAAATGCTGTACGTGGACGTGCAGATCAAGGCTTT	960
Qy	961	GGCGTAGTACCCGAAGGGAACTGAAAGCTGGCCGTGGCGGCAAGGCGCGTAAACGTGGTAA	1020
Db	961	GGCGTAGTACCCGAAGGGAACTGAAAGCTGGCCGTGGCGGCAAGGCGCGTAAACGTGGTAA	1020
Qy	1021	AAAGTGTGATGACCAACGGGTGTCTTTGACATCTCGACACCGGGCAAGTCTTATCTG	1080
Db	1021	AAAGTGTGATGACCAACGGGTGTCTTTGACATCTCGACACCGGGCAAGTCTTATCTG	1080
Qy	1081	GCAAAATGCCCGCAAGCTGGGTGACCGCTTGATTTGTGGCGTCAACAGCGATGECTCCACC	1140
Db	1081	GCAAAATGCCCGCAAGCTGGGTGACCGCTTGATTTGTGGCGTCAACAGCGATGECTCCACC	1140
Qy	1141	AAACGGCTGAAAGGGGATTTCCGCCCGGTAAACCACTGCAACAGCTATGATTTGTCTG	1200
Db	1141	AAACGGCTGAAAGGGGATTTCCGCCCGGTAAACCACTGCAACAGCTATGATTTGTCTG	1200
Qy	1201	GGGCACTGGAAAGCGGTGCACTGGGTAGTGTGTTTGAAGGACAGCGCGACGGCTTG	1260
Db	1201	GGGCACTGGAAAGCGGTGCACTGGGTAGTGTGTTTGAAGGACAGCGCGACGGCTTG	1260
Qy	1261	ATCGCCGGGATCTTGTCCAGATCTGTGTGTAAAGCGCGCACTATAAACGAGAAGATTT	1320
Db	1261	ATCGCCGGGATCTTGTCCAGATCTGTGTGTAAAGCGCGCACTATAAACGAGAAGATTT	1320
Qy	1321	GCCGGGAGTAAAGAAAGTCTGGGCCAACCGGTGGCGAAGTGTGGTGTCTCAACTTTGAAGAC	1380
Db	1321	GCCGGGAGTAAAGAAAGTCTGGGCCAACCGGTGGCGAAGTGTGGTGTCTCAACTTTGAAGAC	1380
Qy	1381	GGTTGTCTCGACGACCAACATCATCAAGAAATCCAAACGAGTATAAAAAGGCTAA	1434
Db	1381	GGTTGTCTCGACGACCAACATCATCAAGAAATCCAAACGAGTATAAAAAGGCTAA	1434

## RESULT 2

AAS52567  
ID AAS52567 standard; DNA; 1434 BP

AC AAS52567;

DT 13-FEB-2002 (first entry)

DE E. coli DNA for cellular proliferation protein #289.

KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

**OS Escherichia coli.**

PN W0200170955-A2.

PD 27-SEP-2001

PF 21-MAR-2001; 2001WO-US009180.

PR 21-MAR-2000; 2000US-0191078P

PR 26-MAY-2000; 2000US-0207727P

PR 27-NOV-2000; 2000US-0253625P

PR 16-FEB-2001; 2001US-0269308P

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlson KL, Zyskind JW, Wall D, Trawick JD, Carr GJ

XX







ID ACA18560 standard; DNA; 1434 BP.  
 AC ACA18560;  
 DT 19-JUN-2003 (first entry)  
 DE Prokaryotic essential gene #217.  
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KM drug design; gene.  
 OS Escherichia coli.  
 PN W020027183-A2.  
 PD 03-OCT-2002.  
 PE 21-MAR-2002; 2002WO-US009107.  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-0072851.  
 PR 06-MAR-2002; 2002US-0362639P.  
 PA (BLIT-) BLITRA PHARM INC.  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 DR WPI; 2003-029926/02.  
 P-PSDB; ABU14690.  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PS Claim 14; SEQ ID NO 6430; 1766bp; English.  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 643 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) providing a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation or to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences  
 CC Sequence 1434 BP; 328 A; 352 C; 447 G; 307 T; 0 U; 0 Other;

Query Match	100.0%; Score 1434; DB 8; Length 1434;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1434;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAGTAAACGCTGCCAGAGTTTGAACGTGACAGAGTGAATGCTGTGATGATG 60	
DB 1 ATGAAAGTAAACGCTGCCAGAGTTTGAACGTGACAGAGTGAATGCTGTGATGATG 60	
QY 61 CTGATGCTTACTGCTGACGCGCCCAACAGTCGTATCTCGCCGGAAGCGCGGCCGTG 120	
DB 61 CTGATGCTTACTGCTGACGCGCCCAACAGTCGTATCTCGCCGGAAGCGCGGCCGTG 120	
QY 121 GTTAAAGTAAATACATCGAAGACGTCGCGGCGCGCGCTACAGTACGTAATATC 180	
DB 121 GTTAAAGTAAATACATCGAAGACGTCGCGGCGCGCGCTACAGTACGTAATATC 180	
QY 181 GCTTCTCTCGGTCTAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240	
DB 181 GCTTCTCTCGGTCTAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240	
QY 241 GCGCTGATGTAATCTCTGCGCGCACTGCAAGTCAATGCAATCTGCTTCTGTAACGAC 300	
DB 241 GCGCTGATGTAATCTCTGCGCGCACTGCAAGTCAATGCAATCTGCTTCTGTAACGAC 300	
QY 301 CATCCGACATTAACAAATTAACGAGTACTTCCCGCAACCAAGCTGATCCGCTGAT 360	
DB 301 CATCCGACATTAACAAATTAACGAGTACTTCCCGCAACCAAGCTGATCCGCTGAT 360	
QY 361 TTTGAAGAGTCTGCAAGGTGTTGATCCGACGCGCTGACGACGAGATTAATCAGCGC 420	
DB 361 TTTGAAGAGTCTGCAAGGTGTTGATCCGACGCGCTGACGACGAGATTAATCAGCGC 420	
QY 421 CTGATGCTGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480	
DB 421 CTGATGCTGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480	
QY 481 CAGCAGATGATCAACCTGCGCGCTGTAAGCGGAGTCTCGGTGCTGATTCGCAAAAGT 540	
DB 481 CAGCAGATGATCAACCTGCGCGCTGTAAGCGGAGTCTCGGTGCTGATTCGCAAAAGT 540	
QY 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAACCGCGAATCTCTCGAATTTGAA 600	
DB 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAACCGCGAATCTCTCGAATTTGAA 600	
QY 601 GCTGTTGCTGCTGTAATGTAAGACCGAAGAAAGATTTGAGCGCGCATGAATCGATT 660	
DB 601 GCTGTTGCTGCTGTAATGTAAGACCGAAGAAAGATTTGAGCGCGCATGAATCGATT 660	
QY 661 GCGGATTTAAGACCTGCGCTGCTGTTAGTACCGGTTCCGACAGAGGTATGCTGCTG 720	
DB 661 GCGGATTTAAGACCTGCGCTGCTGTTAGTACCGGTTCCGACAGAGGTATGCTGCTG 720	
QY 721 CAACCGGTTAAGCGCGCGCTGATATGCAACCAAGCGGAGATGTAAGCTTACC 780	
DB 721 CAACCGGTTAAGCGCGCGCTGATATGCAACCAAGCGGAGATGTAAGCTTACC 780	
QY 781 GGTGCGGCGCAACCGGTATGCGCTGCTGCGCGCAACGCTGCAACGCGGTATTTGCTG 840	
DB 781 GGTGCGGCGCAACCGGTATGCGCTGCTGCGCGCAACGCTGCAACGCGGTATTTGCTG 840	
QY 841 GAAGAAAGCTGCTTCTTGGCAATGCGGCGCGCTGCGGTGCTGCGCAACCTGCGAACC 900	
DB 841 GAAGAAAGCTGCTTCTTGGCAATGCGGCGCGCTGCGGTGCTGCGCAACCTGCGAACC 900	
QY 901 TCCACGCTTTCGCGATGCAAGCTGGAATATGCTGATGCTGATGATGATGATGATGAT 960	
DB 901 TCCACGCTTTCGCGATGCAAGCTGGAATATGCTGATGCTGATGATGATGATGATGAT 960	
QY 961 GCGGTGATGACCAAGAAAGTGAAGCTGCGCTGCGCGCGCGCTGCAACCTGCTGAA 1020	
DB 961 GCGGTGATGACCAAGAAAGTGAAGCTGCGCTGCGCGCGCGCTGCAACCTGCTGAA 1020	
QY 1021 AAGGTGTATGACCAACGCTGCTTGTGATCTCTGACAGCGCGGACGCTCTTATCTG 1080	



Db 1021 AAGTGGATGATGACCAACGAGTGTCTTGACATCTCGACCGCGGACGCTCTTATCTG 1080  
 Qy 1081 GCATATGCCCCGACGAGCTGGGTGACCGCTTGATGTTGTCGCCGACACAGCATGCTCCACC 1140  
 Db 1081 GCATATGCCCCGACGAGCTGGGTGACCGCTTGATGTTGTCGCCGACACAGCATGCTCCACC 1140  
 Qy 1141 AAGCGCTGAAAGGGGATTCGCCGCCGCTAAACCACTCGAACAGCGATGATGCTG 1200  
 Db 1141 AAGCGCTGAAAGGGGATTCGCCGCCGCTAAACCACTCGAACAGCGATGATGCTG 1200  
 Qy 1201 GCGGCACCTGAAAGCGCTGACCTGAGTGTGTTGAAAGAGACACGCGGACGCTTG 1260  
 Db 1201 GCGGCACCTGAAAGCGCTGACCTGAGTGTGTTGAAAGAGACACGCGGACGCTTG 1260  
 Qy 1261 ATCCCGCGGATCTTCCAGATCTGCTGTGAAAGCGGCGCATTTAAACCGAAGAGATT 1320  
 Db 1261 ATCCCGCGGATCTTCCAGATCTGCTGTGAAAGCGGCGCATTTAAACCGAAGAGATT 1320  
 Qy 1321 GCCCGAGTAAAGAGTCTGGGCGCAACGCTGCGAAGTGTGCTCAACTTTGAAGAC 1380  
 Db 1321 GCCCGAGTAAAGAGTCTGGGCGCAACGCTGCGAAGTGTGCTCAACTTTGAAGAC 1380  
 Qy 1381 GGTGCTCGACGACCAATCATCAAGAGATCCAAACGATTAAGGCTAA 1434  
 Db 1381 GGTGCTCGACGACCAATCATCAAGAGATCCAAACGATTAAGGCTAA 1434

RESULT 5  
 ADH88994  
 ID ADH88994 standard; DNA; 1434 BP.

AC ADH88994;

DT 22-APR-2004 (first entry)

DE Escherichia coli K1 S26 mutant coding sequence, SEQ ID 28.

XX Pathogenic microorganism; K1 S26; mutant; pathogenicity; virulence;  
 KW bacterial infection; extra-intestinal infection; enterobacterium; gene;  
 db.

OS Escherichia coli.

OS Synthetic.

XX Key location/Qualifiers  
 FT CDS 1..1434  
 FT /\*tag= a  
 FT /product= "K1 S26 mutant"

XX W02004005535-A2.

XX 15-JAN-2004.

XX 09-JUL-2003; 2003WO-EP008209.

XX 09-JUL-2002; 2002FR-0008636.

XX (MUT-.) MUTABILIS SA.

XX Becalch S;

XX MPI; 2004-091381/09.

XX P-PSDB; ADH88994.

PT Identifying and selecting a gene required for the proliferation in vivo  
 PT of a pathogenic microorganism comprises determining the virulence of  
 PT mutant genes on an experimental model of infection, and their effect on  
 PT enteric colonization.

PS Claim 5; SEQ ID NO 28; 60BP; English.

XX The present invention relates to a method for identifying and selecting a

CC gene required for the proliferation in vivo of a pathogenic  
 CC microorganism. The method comprises determining the virulence of mutant  
 CC genes on an experimental model of infection, and their effect on enteric  
 CC colonization in an axenic mouse model. ADH88967-ADH88981 and ADH88982-  
 CC ADH88996 are pathogenicity or virulence targets and their coding  
 CC sequences which were used in the method of the invention. Compounds  
 CC capable of inhibiting pathogenicity or virulence target expression are  
 CC useful for developing medicinal products for inhibiting a bacterial  
 CC infection, in particular an extra-intestinal infection in the case of  
 CC enterobacteria.

XX Sequence 1434 BP; 328 A; 352 C; 447 G; 307 T; 0 U; 0 Other;

Qy Query Match 100.0%; Score 1434; DB 12; Length 1434;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAGTAAACGCTGCGACGATTTGAACGTGCAGAGTATGCTGTTGATGATG 60  
 Db 1 ATGAAAGTAAACGCTGCGACGATTTGAACGTGCAGAGTATGCTGTTGATGATG 60  
 Qy 61 CTGGATCTTACCTGATACGCGCCCAACAGCTGATCTGCTGCGAAGCGCGGCGCTG 120  
 Db 61 CTGGATCTTACCTGATACGCGCCCAACAGCTGATCTGCTGCGAAGCGCGGCGCTG 120  
 Qy 121 GTTAAAGTAAATACATCGAAGAACGCTCGCGCGCGCTTAACGTGCGATGAATATC 180  
 Db 121 GTTAAAGTAAATACATCGAAGAACGCTCGCGCGCGCTTAACGTGCGATGAATATC 180  
 Qy 181 GCTTCTCTCGGTCTAATGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Db 181 GCTTCTCTCGGTCTAATGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Qy 241 GCGCTGATTAATCTCTGCGCGCGACGCTCAACGTAATGCGACTTCGTTCTGACGACG 300  
 Db 241 GCGCTGATTAATCTCTGCGCGCGACGCTCAACGTAATGCGACTTCGTTCTGACGACG 300  
 Qy 301 CATCCGACATTAACCAATTAACGCTGATCTTCCGCAACCAACAGCTGATCGCTGAT 360  
 Db 301 CATCCGACATTAACCAATTAACGCTGATCTTCCGCAACCAACAGCTGATCGCTGAT 360  
 Qy 361 TTGAAAGAGGTTTCCAGAGTGTGATCCGACGCGCTGACAGAGGATTAATCAGGCG 420  
 Db 361 TTGAAAGAGGTTTCCAGAGTGTGATCCGACGCGCTGACAGAGGATTAATCAGGCG 420  
 Qy 421 CTGAGTTGCAATTGCGCGCGCTGCTTCTGACTGACGCAAGGTTGCTGCGCAAGCTGA 480  
 Db 421 CTGAGTTGCAATTGCGCGCGCTGCTTCTGACTGACGCAAGGTTGCTGCGCAAGCTGA 480  
 Qy 481 CAGCAGATGATCCCACTGCGCGCTGAAGCGGATGTTCCGCTGCTGATTCGCAAAAGGT 540  
 Db 481 CAGCAGATGATCCCACTGCGCGCTGAAGCGGATGTTCCGCTGCTGATTCGCAAAAGGT 540  
 Qy 541 ACCGATTTTGAAGCGCTACCGCGCGCTGCTGCTTAAACCGGAAATCTCTGGAATTTGAA 600  
 Db 541 ACCGATTTTGAAGCGCTACCGCGCGCTGCTGCTTAAACCGGAAATCTCTGGAATTTGAA 600  
 Qy 601 GCTGTTGCTGTAATGTAAGACCGAAGAGATGTTGAAGCGCGCATGAATGAT 660  
 Db 601 GCTGTTGCTGTAATGTAAGACCGAAGAGATGTTGAAGCGCGCATGAATGAT 660  
 Qy 661 GCCGATTAACGACTCGGCTGTTAGTACCGGTTCCGACAGGGTATGCTGCTG 720  
 Db 661 GCCGATTAACGACTCGGCTGTTAGTACCGGTTCCGACAGGGTATGCTGCTG 720  
 Qy 721 CAACCGGTTAAAGCGCGCTGCAATATGCCAACCAAGCGAGAGTATGACGTTAAC 780  
 Db 721 CAACCGGTTAAAGCGCGCTGCAATATGCCAACCAAGCGAGAGTATGACGTTAAC 780  
 Qy 781 GGTGCGGCGGACAGCGGATGAGCGCTGCGCGGCAACGCTGACAGCGGATTAATGCTG 840  
 Db 781 GGTGCGGCGGACAGCGGATGAGCGCTGCGCGGCAACGCTGACAGCGGATTAATGCTG 840

```

Oy 841 GAAGAAGCCTGCTTCTTCCCAATGCGCGGCTGAGTGTGTGCGCAAACTGGGAACC 900
Db 841 GAAGAAGCCTGCTTCTTCCCAATGCGCGGCTGAGTGTGTGCGCAAACTGGGAACC 900
Oy 901 TCACGAGTTTCCGCGATGAGCTGGAATAATGCTGTACAGTGCAGATACAGGCTTT 960
Db 901 TCACGAGTTTCCGCGATGAGCTGGAATAATGCTGTACAGTGCAGATACAGGCTTT 960
Oy 961 GGCCTGATGACCGGAAGAACTGAAGCTGCGCCGTAAGCGGCGCTTAACTGTGTAA 1020
Db 961 GGCCTGATGACCGGAAGAACTGAAGCTGCGCCGTAAGCGGCGCTTAACTGTGTAA 1020
Oy 1021 AAAGTGTGATGACCAACGATGCTTGTATCATCTGTACAGCGCGGCACTCTTATCTG 1080
Db 1021 AAAGTGTGATGACCAACGATGCTTGTATCATCTGTACAGCGCGGCACTCTTATCTG 1080
Oy 1081 GCAAAATGCGCGCAAGCTGGGTGACCGCTTGAATGTTGCGTCAACAGCGATGCTTCAAC 1140
Db 1081 GCAAAATGCGCGCAAGCTGGGTGACCGCTTGAATGTTGCGTCAACAGCGATGCTTCAAC 1140
Oy 1141 AAACGGCTGAAAAGGGGATTTCCCGCCCGTAAACCACTGCAACAGCGTATGATGCTG 1200
Db 1141 AAACGGCTGAAAAGGGGATTTCCCGCCCGTAAACCACTGCAACAGCGTATGATGCTG 1200
Oy 1201 GACGCACTGAAAAGCGGTGATGCTGTGTGAAAGAGACAGCGCGAGCGCTTG 1260
Db 1201 GACGCACTGAAAAGCGGTGATGCTGTGTGAAAGAGACAGCGCGAGCGCTTG 1260
Oy 1261 ATCCGCGGATCTTGCCAGATCTGCTGTGAAAGCGCGCACTATTAACAGAAAGATT 1320
Db 1261 ATCCGCGGATCTTGCCAGATCTGCTGTGAAAGCGCGCACTATTAACAGAAAGATT 1320
Oy 1321 GCCCGAGTAAAGAGTCTGCGCCCAACGCTGCGCAAGTGTGTCTCACTTTGAAGAC 1380
Db 1321 GCCCGAGTAAAGAGTCTGCGCCCAACGCTGCGCAAGTGTGTCTCACTTTGAAGAC 1380
Oy 1381 GGTGTCTGACGACCAATCATCATCAAGAGATCCAAAGATTAAGGCTTAA 1434
Db 1381 GGTGTCTGACGACCAATCATCATCAAGAGATCCAAAGATTAAGGCTTAA 1434

```

## RESULT 6

AC32148 standard; DNA; 1428 BP.

```

ID AC32148 standard; DNA; 1428 BP.
AC AC32148,
XX 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #13805.
XX KW Antisense; dr, prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Enterobacter cloacae.
XX PN W020027183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (EliT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW,
XX PI Mail D, Travick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

```

```

DR WPI; 2003-029926/02.
DR P-PSDB; ABU28278.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 20018; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1428 BP; 292 A; 388 C; 489 G; 259 T; 0 U; 0 Other;

```

Query Match 71.0%; Score 1018.2; DB 8; Length 1428;  
 Best Local Similarity 82.2%; Pred. No. 8.5e-282;  
 Matches 1170; Conservative 0; Mismatches 253; Indels 0; Gaps 0;

```

Oy 1 ATGAAAGTAAAGCTGCCAGAGTTTGAACGTGACGAGTGTGTGTGATGTGATG 60
Db 1 ATGAAAGTAAAGCTGCCAGAGTTTGAACGTGACGAGTGTGTGTGATGTGATG 60
Oy 61 CTGATCTTACTGTGATGCGGCCCCACCAAGTGTATCTGCGCGAAGCGCGTGCCTG 120
Db 61 CTGATCTTACTGTGATGCGGCCCCACCAAGTGTATCTGCGCGAAGCGCGTGCCTG 120
Oy 121 GTTAAAGTAAATACATCGAAGAAAGTCCGGGGCGCGCTTAACCTGGAGTAATTC 180
Db 121 GTTAAAGTAAATACATCGAAGAAAGTCCGGGGCGCGCTTAACCTGGAGTAATTC 180
Oy 181 GCTTCTCTGCGGTCTAATGACCGCTGTGTGAGTGAAGCGGCAATGACATGACGCGC 240
Db 181 GCTTCTCTGCGGTCTAATGACCGCTGTGTGAGTGAAGCGGCAATGACATGACGCGC 240
Oy 241 GCGCTGATGAATCTCTGCGCGACGTCAACGTCAATGCGACTTCTGTGTACGACG 300
Db 241 GCGCTGATGAATCTCTGCGCGACGTCAACGTCAATGCGACTTCTGTGTACGACG 300
Oy 301 CATCCGACATTAACAAATTAACGGTACTTCCCGCAACCAAGAGCTGATCGAT 360
Db 301 CATCCGACATTAACAAATTAACGGTACTTCCCGCAACCAAGAGCTGATCGAT 360
Oy 361 TTGGAAGAGTGTGCAAGGTGTGATCCGACCGGTGACGACGAGATTAATCAAGCG 420

```

```
Dh 361 TTTGAGAGAGATTGAGAGCGTTGATCCGAGCGCTGACGAGCGTATCAACGACG 420
Qy 421 CTGAGTTGATTTGGCGGCTGCTGCTTCTGACTACGCAAGGAGGCTGGCAAGGTA 480
Dh 421 CTGGCAATATTTGGCGGCTGCTGCTTCTGACTACGCAAGGAGGCTGGCAAGGTA 480
Qy 481 CAGCAGATGATCAACTGGCGCGTAAAGCGGGTCTCCGCTGCTGATTTGCAAAAGT 540
Dh 481 AAGAGATGATTTCAAGCTGGCGGCTAAAGCAACGCTCCGGTGTGATGACCCGAAAGG 540
Qy 541 ACCGATTTTGAAGCGCTACCGCGGGCTACGCTTTTAAAGCGGATCTCTCGAATTTGA 600
Dh 541 ACCGATTTTGAAGCGCTACCGCGGGCTACGCTTTTAAAGCGGATCTCTCGAATTTGA 600
Qy 601 GCTGTGTGCGTAAATGATGAGACGGAAGAGATTTGAGCGCGGATGAACTGAT 660
Dh 601 GCTGTGTGCGGCAAGTGCAGAAACGAGAGAGAGCTGTTGAGCGCGGATGAACTGAT 660
Qy 661 GCCGATTAAGAACTCTCGGCTCTGTGATGACCGGTTCCGAAAGGATGTCGCTG 720
Dh 661 GCCGATTTTGAAGCTGCTGCGGCTGCTGATGACCGGCTCCGAGAGGAGATGACGCTG 720
Qy 721 CAACCGGATTAAGCGCGCTGCTGATTTGCAACCCAGCGCGAGAGTGTATGACGTTACC 780
Dh 721 CAACCGGCGCAAGCGCGCTGCTGATTTGCAACCCAGCGCGAGAGTGTATGACGTTACC 780
Qy 781 GGTGCGGCGCGACGAGTATTTGGCGCTCTGCGCGCAACGCTGAGCGGATTTGCTG 840
Dh 781 GGTGCGGCGCGACGAGTATTTGGCGCTCTGCGCGCAACGCTGAGCGGATTTGCTG 840
Qy 841 GAAGAGAGCTGCTCTTTGCAATGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 900
Dh 841 GAAGAGAGCTGCTCTTTGCAATGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 TCCAGCTTTTGGCGGCTGCTGAGCTGAGAAATCTGTAGCTGAGACGCTGAGATACGCTT 960
Dh 901 TCAACCGTTTTCGCAATCTGAGCTGAGAAACCGGCTGCGCTGCTGCTGCTGCTGCTT 960
Qy 961 GGCCTGATGACGAGAGAGAACTGAAAGCTGAGCGCGGAGCGCGCTGAAACGCTGTA 1020
Dh 961 GGCCTGATGACGAGAGAGAACTGAAAGCTGAGCGCGGAGCGCGCTGAAACGCTGTA 1020
Qy 1021 AAAGTGTGATGACCAACGCTGCTTTGACATCTTGCACGCGCGGAGCAAGCTCTTATCTG 1080
Dh 1021 AAAGTGTGATGACCAACGCTGCTTTGACATCTTGCACGCGCGGAGCAAGCTCTTATCTG 1080
Qy 1081 GCAAAATGCCCGCAAGCTGGGCTGACCGCTTGAATTTGCTGCTGCAACGAGATGCTTCAAC 1140
Dh 1081 GCGAAGCGCGCAAGCTGGGCTGACCGCTTGAATTTGCTGCTGCAACGAGATGCTTCAAC 1140
Qy 1141 AAAGCGCTGAAAGGAGGATTTCCCGCCCGGTAAACCACTCGAAGACGCTATGATGCTG 1200
Dh 1141 AAGCGCTGAAAGGAGGAGGATTTCCCGCCCGGTAAACCACTCGAAGACGCTATGATGCTG 1200
Qy 1201 GCGCACTGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Dh 1201 GGTGCGCTGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Qy 1261 ATCCCGCGGATCTTGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Dh 1261 ATTCGCGGATCTTGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Qy 1321 GCGCGAGTAAAGAGTCTGGGCGCAACGCTGGCGAGAGTGTGTCTCAACTTTGAAGAC 1380
Dh 1321 GCGCGAGTAAAGAGTCTGGGCGCAACGCTGGCGAGAGTGTGTCTCAACTTTGAAGAC 1380
Qy 1381 GGTGCTGACGACCAATCATCAAGAGATCCAAAGATTA 1423
Dh 1381 GGTGCTGACGACCAATCATCAAGAGATCCAAAGATTA 1423
```

RESULT 7  
ACH96242

```
ID ACH96242 standard; DNA; 1500 BP.
XX
AC ACH96242:
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polynucleotide seqid 2037.
XX
KW Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; de.
XX
OS Klebsiella pneumoniae.
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
PE 27-JAN-2000; 2000US-00489039.
XX
PR 29-JAN-1999; 99US-0117747P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL, Osborne M;
XX
DR WPI; 2003-895346/82.
XX
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS Disclosure; SEQ ID NO 2037; 932pp; English.
XX
CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention
XX
SQ Sequence 1500 BP; 308 A; 416 C; 497 G; 279 T; 0 U; 0 Other;
XX
Query Match 70.6%; Score 1011.8; DB 11; Length 1500;
Best Local Similarity 81.9%; Pred. No. 7e-290;
Matches 1166; Conservative 0; Mismatches 257; Indels 0; Gaps 0;
Qy 1 ATGAAAGTAAAGCTGCCAGAGTTTGAACGTGACGAGATGATGCTGCTGCTGCTGCTGCTG 60
Dh 67 ATGAAAGTAAAGCTGCCAGAGTTTGAACGTGACGAGATGATGCTGCTGCTGCTGCTGCTG 126
Qy 61 CTGAGCTGTAATCTGTAAGGCGCCCAACGATGCTATCTGCGCGGAGCGCGGCTGCTGCTG 120
Dh 127 CTGAGCTGTAATCTGTAAGGCGCCCAACGATGCTATCTGCGCGGAGCGCGGCTGCTGCTG 186
Qy 121 GTTAAAGTAAATCAATGAAAGAGCTCGGCGCGCGCTTAAAGTGAATGATATC 180
Dh 187 GTGAGGTGAAATATCGAAGAACGCTCGGCGCGCGCTTAAAGTGAATGATATC 246
Qy 181 GCTTCTCTCGCTGCTAATGACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Dh 247 GCTCTGCTGCGGCAACCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
Qy 241 GCGCTGAGTAATCTCTGCGCGAGCTCAAGTCAAAATGCACTTCTGTTTCTGTAACGAG 300
Dh 307 GCGCTGAGCAAGCGCTGCGCAACGCTCAATGTAAGTGCATCTTCTGCTCCGACT 366
Qy 301 CATCCGACATTAACCAATTAACGCTATCTTCCGCAACCAAGCTGATCCGCTGAT 360
Dh 367 CACCGGATATGACCAAGAGTGGGGTGTGTGCGGAAATCAAGAGTATCCGCTGAC 426
Qy 361 TTTGAAGAGGTTTGAAGGTTGATCCGACCGCTGACGAGCGGATTAATCAAGCG 420
```

```

Db 427 TTGAGAGAGGCTTCTCCGCGCTGATCCGACGCGATGACAGAGCGATTGACAGGCG 486
Qy 421 CTGAGTTGATTTGGCGCGCTGCTTCTTCTGATACGCAAAAGTGGCGTGGCAAGGTA 480
Db 487 CTGGGCTTATTTGGCGCGCTGGTCTTTCGATATGCGCAAAAGGCGGCTGACAGCGGTC 546
Qy 481 CAGCAGATGATCCAACTGCGCGCTGAAAGCGGCTGTTCCGCGTGTGATTTGATCCAAAAGT 540
Db 547 CAGACATGATCCGATCTGACGATGAAAGCGGCGTGGCTGCTGATGATCCGAAAGGC 606
Qy 541 ACCGATTTGAGCGCTACCGCGCGCTACCGCTTTAAACCGCAATCTCTCGGAAATTTGAA 600
Db 607 ACCGATTTGAGCGCTATGCGCGCGCTGCTGATCCGCGAATCTCTCTGAGTTTGA 666
Qy 601 GCTGTTTCCGTAATGTAAGACCGAAAGAAAGATTTGAGCGCGCTGTAATCTGATT 660
Db 667 GCGGTGATGCGCAAGTCCAGATGAAAGCGGATGCTGTAAGCGCGGCAATGAACTCATC 726
Qy 661 GCCGATTAAGAACTCTGGGCTCTGTTAGTACCGCGTCCGAAAGGATGATGCGTCTG 720
Db 727 GCCGATTAAGAACTCTGGGCTCTGTTAGTACCGCGTCCGAAAGGATGATGCGTCTG 786
Qy 721 CAACCGGATTAAGCGCGCTGCTGATATGCAACCGCAAGCGCAAGATGATGACGTTACC 780
Db 787 CAGCGGGAAGTCCGCGCTGCTATATGCGCAACCGCAAGCGCAAGATGATGACGTTACC 846
Qy 781 GGTGCGGCGCAACAGGATGATTTGGGCTCTGCGCGCAACGCTGCGAGCGGTAATTCGCTG 840
Db 847 GCGCGCGCGCAACAGGATGATTTGGGCTCTGCGCGCAACGCTGCGAGCAACCTG 906
Qy 841 GAAGAAAGCTGCTCTTTGCAATGCGGCGGCTGCGGCTGATGATGCGCAACCTG 900
Db 907 GAAGAAAGCTGCTCTTTGCAATGCGGCGGCTGCGGCTGATGATGCGCAACCTG 966
Qy 901 TCACCGGCTTTCGCGGATGAGCTGAAATGCTGATGCGGCAACCTG 960
Db 967 TCACCGGCTTTCGCGGATGAGCTGAAATGCTGATGCGGCAACCTG 1026
Qy 961 GCGCTGATGACCGAAAGGAACTGAAAGTGGCGCTTACCGGCGGCTTAAACGTTGGTGA 1020
Db 1027 GCGCTGATGACCGAAAGGAACTGAAAGTGGCGCTTACCGGCGGCTTAAACGTTGGTGA 1086
Qy 1021 AAATGATGATGACCAACGCGTGTCTTTGACATCTCTGACGCGCGGCAAGTCTTTATCTG 1080
Db 1087 AAATGATGATGACCAACGCGTGTCTTTGACATCTCTGACGCGCGGCAAGTCTTTATCTG 1146
Qy 1081 GCAAAATCCCGCAAGCTGCGTGAACCGCTTGAATTTGCGCGTCAACAGCGATGCTTCAAC 1140
Db 1147 GCGAATCCCGCAAGCTGCGTGAACCGCTTGAATTTGCGCGTCAACAGCGATGCTTCAAC 1206
Qy 1141 AAATGATGATGACCAACGCGTGTCTTTGACATCTCTGACGCGCGGCAAGTCTTTATCTG 1200
Db 1207 AAATGATGATGACCAACGCGTGTCTTTGACATCTCTGACGCGCGGCAAGTCTTTATCTG 1266
Qy 1201 GCGGCACTGAAAGGCTGCACTGCGTGTGCTTTGAAAGGACAGCGCGGCGCTG 1260
Db 1267 GCGGCACTGAAAGGCTGCACTGCGTGTGCTTTGAAAGGACAGCGCGGCGCTG 1326
Qy 1261 ATGCGCGGATCTTTCGAGATCTGCTGTTGAAAGCGCGCACTTAAACGAGAGATTT 1320
Db 1327 ATGCGCGGATCTTTCGAGATCTGCTGTTGAAAGCGCGCACTTAAACGAGAGATTT 1386
Qy 1321 GCGGCACTGAAAGGCTGCACTGCGTGTGCTTTGAAAGGACAGCGCGGCGCTG 1380
Db 1387 GCGGCACTGAAAGGCTGCACTGCGTGTGCTTTGAAAGGACAGCGCGGCGCTG 1446
Qy 1381 GGTGCTGACGACCAATCATCAAGAAAGATCCAGAGATTA 1423
Db 1447 GGTGCTGACGACCAATCATCAAGAAAGATTA 1489

```

```

ID AAS56024 standard; DNA; 1434 BP.
XX
AC AAS56024;
XX
DT 13-FEB-2002 (first entry)
XX
DE Salmonella typhi DNA for cellular proliferation protein #57.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX antibiotic; drug design.
XX
OS Salmonella typhi.
XX
PN M0200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001MO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (BLIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
XX
DR P-PDB; A038165.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX
PS Claim 27; SEQ ID NO 9661; 511bp; English.
XX
XS
XS The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Streptococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 1434 BP; 311 A; 360 C; 468 G; 295 T; 0 U; 0 Other;

```

```

Query Match 70.3%; Score 1008.6; DB 4; Length 1434;
Best Local Similarity 81.8%; Pred. No. 6.1e-289;
Matches 1164; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

```

```

Qy 1 ATGAAGTAAAGCTGCGCAGAGTTGAAAGTGCAGAGTATGCTGGTGGTATGATG 60
Db 1 ATGAAGTAAAGTTCGCAAGGTTTGAACGTGACGCGCTATGTTGGTATGATG 60
Qy 61 CTGATGCTTACTGATGCGCGCCGACAGTGTATCTCGCGGAGCGCGGCTG 120
Db 61 CTGATGCTTATGATGCGCGCCGACAGTGTATTCACCGGAGCGCGGCTG 120

```

QY 121 GTTAAAGTAATACATCGAAGAACTCCGGGCGCGCGCTAAAGTGGCGATGAAATATC 180  
 Db 121 GTTAAAGTAATACATCGTGAAGAAAGTCCGGGCGCGCGGAAAGTGGCGATGAAATATC 180  
 QY 181 GCTTCTCTCGGTGCTAATGACAGCCTGGTGGGTTGACGGGCAATTTGACATGACGGCCG 240  
 Db 181 GCGTCTCTGGGAGGAAACGCCGCTGCTGGTCCGCTGACGGGTATTTGACGCCGCCGCG 240  
 QY 241 GCGGTAGTAAATCTGCGGCGGAGCTGCAAGTCAATGCGCACTTGGTTTGTACCGAGC 300  
 Db 241 GCGGTAGTAAATCTGCGGCGGAGCTGCAATGCAATGCGCACTTGGTTTGTACCGAGC 300  
 QY 301 CATCGACCATTAACCAATTAACGGGATCTTTCCGCAACCAACAGCTGATCCGTGAT 360  
 Db 301 CATCGACCATTAACCAATCTGGGTACTATCAGTAAATGACAGCTCATCTCGTCTTAT 360  
 QY 361 TTTGAAGAAGGTTTCGAAGGTGTTGATCCGACGCCCTGACAGCGGATTAATCAGGCG 420  
 Db 361 TTTGAAGAAGGCTTTGAGGGGCTGGAACCCGACGCCCTTGGCATGACGCTATCAGCAG 420  
 QY 421 CTGAGTTGATTTGGCGCGCTGGTCTTCTGACTACGCGCAAAAGGTCGCGCAAGCGTA 480  
 Db 421 CTGAGTTGATTTGGCGCGCTGGTCTTCTGACTACGCGCAAAAGGTCGCGCAAGCGTA 480  
 QY 481 CAGCAGATGATCAACTGGCGCGTAAAGCGGCTGTTCCGCTGCTGATTTGATCCAAAGT 540  
 Db 481 CAGCAGATGATTTCCCTAGCGCGCCAGCGCGGCGCTGCGTCTCATCGATCCGAAAGGA 540  
 QY 541 ACCGATTTTGAAGCGCTACCGCGGCGCTACGCTGTTAACGCCGAATCTCTCGGAAATTTGAA 600  
 Db 541 ACCGATTTTGAAGCGCTTACCGCGGCGCAACGCTGCTGACGCCAAATCTTTCTGAATTTGAG 600  
 QY 601 GCTGTTGCGGTAAATGTAAGACCCGAAAGAGAATGTTGTTAGCGCGGCAATGAATCTGAT 660  
 Db 601 GCGGTTGCGGGAATGTAAGACCCGAAAGAGAATGTTGTTAGCCGCGATTAACCTGAT 660  
 QY 661 GCGGATTAAGAACTCTCGGCTCTGTTAGTGAACCGGTTCCGAAAGGTTAGTCTGCTG 720  
 Db 661 GCGGATTAAGAACTCTTTCGCGCTGTTGGTCAACGCTTCCGAAAGGTTAGTCTGCTG 720  
 QY 721 CAACCGGTTAAAGGCGGCTGCTGATGTCGCAACCGAGCGGAGAGTGTATGAGCTTACC 780  
 Db 721 CAACCGAATTAAGGCGGCTGCTGATGTCGCAACCGAGCGGAGAGTGTATGAGTATACC 780  
 QY 781 GGTGCGGCGGACACGAGTATGCGCTCTGCGCGGCAACGCTGAGCGGTAATTCGCTG 840  
 Db 781 GGTGCGGCGGACATCGATGATCGGCTGCTGCGCGGCAACGCTGAGCGGTAATTCGCTG 840  
 QY 841 GAAGAAGCTCTCTTCTTGGCAATGCGGCGGCTGCGCTGCTGCTGCGCAATCTGCGAACC 900  
 Db 841 GAGGAGCGGTGTATTTTGGCAATGCGGCGGCGGCTGCTGCTGCTGCGCAATCTGCGGACG 900  
 QY 901 TCCACGCTTTGCGGATGCACTGGAATAATCTGTACTGGAACGTGCAATACAGGCTTT 960  
 Db 901 TCCACGCTTTGCGGATGCACTGGAATAATCTGTACTGGAACGTGCAATACAGGCTTT 960  
 QY 961 GGCCTGATGACCGAAGAGAACTGAAAGTGGCGGCTGAGCGGCGCGTAAACGTGTTGAA 1020  
 Db 961 GGCCTGATGACCGAAGAGAGAGTGAAGAGCGCGCTGAGCGGCGCGTAAACGTGTTGAG 1020  
 QY 1021 AAAGTGTGATGACCAACGCTGTCTTTGACATCTGCAACGCGCGGACGCTCTTTATCTG 1080  
 Db 1021 AAAGTGTGATGACCAACGCGGCTTTGATATTTGCAACGCGCGGACGCTCTTTATCTG 1080  
 QY 1081 GCAAAATGCCCGCAAGCTGGTGAACGCTGATTTGTCGCTGCAACAGCGATGCTTCAACC 1140  
 Db 1081 GCGAAGCGCGCAACCTGGGCAACGCTGATTTGTCGCTGCAACAGCGATGCTTCAACT 1140  
 QY 1141 AAAGCGTGAAGAGGAGATTTCCGCGCGGTTAAACCACTCGAAGACGCTGATTTGCTG 1200  
 Db 1141 AAAGCTGGAAGAGGAGAGCGCTGCGGTTAAATCCGCTCGAAGACGCTGATTTGCTG 1200  
 QY 1201 GCGCGACTGGAAGCGGCTGCACTGGGTAGTGTCTTTGAAAGAGACACGCCGACGCTTG 1260

Db 1201 GCGCGCGTGAAGTCCGCTGCACTGGTGTCTCTTTGAAAGAGGATACGCCGCAACGACTG 1260  
 QY 1261 ATGCGCGGATCTTGGCAATCTGCTGATGTAAGGCGGCGGCAATTAACGAGAAAGATT 1320  
 Db 1261 ATGCGGATTTCTGCGGATCTGCTGATGTAAGGCGGCGGCAATTAAGCTGGAAGATC 1320  
 QY 1321 GCGGAGTAAAGATCTGGGCAACGCTGCGGAGAGTGTGCTCACTTTGAAGAC 1380  
 Db 1321 GCGGAGCAAGAAAGGCTTGGGCAACGCGGCGGAGATGATGCTGCAACTTGAAGAT 1380  
 QY 1381 GGTGCTCGACGACCAATCATCAAGAGATCAACAGATA 1423  
 Db 1381 GGTGCTCGACGACCAATATCATCAAAAGATCAACAGAGA 1423  
 RESULT 9  
 ACAS1424  
 ID ACAS1424 standard; DNA; 1434 BP.  
 AC  
 ACAS1424;  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Prokaryotic essential gene #33081.  
 DE Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX Salmomella typhl.  
 OS  
 PN W0200277183-A2.  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002MO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen Ku, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR P-PSDB; ABU47554.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids, required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PS  
 PS Claim 14; SBQ ID NO 39294; 1766pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 613 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the pinned specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pat\\_sequences](http://wipo.int/pub/published_pat_sequences)

XX Sequence 1434 BP; 311 A; 360 C; 468 G; 295 T; 0 U; 0 Other;

Query Match 70.3%; Score 1008.6; DB 8; Length 1434;  
Best Local Similarity 81.8%; Pred. No. 6.1e-289;  
Matches 1164; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

QY 1 ATGAAGTAAAGCTGCGAGATTGAACTGTCAGAGTATGTGTGTGATGATG 60  
DB 1 ATGAAGTAAATCTGCCAGCTTTGAACTGTCAGAGCTATGTGTGTGATGATG 60  
QY 61 CTGATGTTTCTGTTACGCGCCCAAGTCGTATCTGCGCGAAGCGCGTCCGCTG 120  
DB 61 CTGATGCTTATTTGTATGCGCCCACTGCGTATTTTCAACGGAAGCGCGTCCGCTG 120  
QY 121 GTTAAAGTAAATACATCGAAGAAAGTCCGCGCGCGCTAATCGTGGCGATGATATC 180  
DB 121 GTTAAAGTAAATACGTTGAGAAAGTCCGCGCGCGCTAATCGTGGCGATGATATC 180  
QY 181 GCTTCTCTCGGTGCTAAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 181 GCGCTCTCGGAGGAGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 241 GCGCTGCTAAATCTCTGCGCGAGCTCAACGTAATGCGAATGCTGCTGCTGCTGCTG 300  
DB 241 GCGCTGAGCAAAAGCTGCGCGAGCTCAATGTAAGTGCAGCTTCTGCTGCTGCTG 300  
QY 301 CATCGACATTAACCAATTAAGGGTATCTTCCGCGAAGCAAGCTGATCCGTGATG 360  
DB 301 CATCGACATTAACCAATTAAGGGTATCTTCCGCGAAGCAAGCTGATCCGTGATG 360  
QY 361 TTGGAAGAGGTTTGAAGGTGTGATCCGCGAGCTGTCAGAGCGATTAATCAGCG 420  
DB 361 TTGGAAGAGGCTTTGAGGGGTGAGCCGCGAGCTGTCAGAGCGATTAATCAGCG 420  
QY 421 CTGAGTTGATGCGCGCTGCTGCTTCTGCTAAGCGCAAGGTGCTGCGAAGCTG 480  
DB 421 CTGAGTTGATGCGCGCTGCTGCTTCTGCTAAGCGCAAGGTGCTGCGAAGCTG 480  
QY 481 CAGAGATGATCCAGCTGCGCGCTGTAAGCGGCTTCCGCTGATGATCCAAAGT 540  
DB 481 CAGAGATGATCCAGCTGCGCGCTGTAAGCGGCTTCCGCTGATGATCCAAAGT 540  
QY 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAAGCGGATCTCTCGAATTTGAA 600  
DB 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAAGCGGATCTCTCGAATTTGAA 600  
QY 601 GCTGTTGCTCGTAAATGTAAGACCGAAGAGATGTTGAGCGCGCGATGAACTGAT 660  
DB 601 GCTGTTGCTCGTAAATGTAAGACCGAAGAGATGTTGAGCGCGCGATGAACTGAT 660  
QY 661 GCGGATTAAGAACTCTGTTAGTGAACCGCTTCCGAAAGCGGATGATGCTGCTG 720  
DB 661 GCGGATTAAGAACTCTGTTAGTGAACCGCTTCCGAAAGCGGATGATGCTGCTG 720  
QY 721 CAACCGGATTAAGCGCGCTGCTGATGTCAGAACCGAAGCGCGAAGTGTATGAGCTTACC 780  
DB 721 CAACCGGATTAAGCGCGCTGCTGATGTCAGAACCGAAGCGCGAAGTGTATGAGCTTACC 780

QY 781 GTCGCGCGCGACACGCGGATTTGCGCTGCTGCGCGCAACGCTGCGAGCGGTAATGCTG 840  
DB 781 GTCGCGCGCGATACGCGGATTTGCGCTGCTGCGCGCAACGCTGCGAGCGGTAATGCTG 840  
QY 841 GAAGAAAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900  
DB 841 GAAGAAAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900  
QY 901 TCAACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960  
DB 901 TCAACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960  
QY 961 GCGGATTAAGAACTCTGTTAGTGAACCGCTTCCGCGCTAATCGTGGCGATGATATC 1020  
DB 961 GCGGATTAAGAACTCTGTTAGTGAACCGCTTCCGCGCTAATCGTGGCGATGATATC 1020  
QY 1021 AAAGTGTGATGACCAACGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080  
DB 1021 AAAGTGTGATGACCAACGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080  
QY 1081 GCAAAATGCGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 1081 GCAAAATGCGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 AAACGCTGTAAGAGGATTTCCCGCGCTAATCCCACTGCAAGCGATGATGCTGCTG 1200  
DB 1141 AAACGCTGTAAGAGGATTTCCCGCGCTAATCCCACTGCAAGCGATGATGCTGCTG 1200  
QY 1201 GCGGCACTGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
DB 1201 GCGGCACTGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
QY 1261 ATGCGCGGATCTTCTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
DB 1261 ATGCGCGGATCTTCTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
QY 1321 GCGGCACTGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
DB 1321 GCGGCACTGGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
QY 1381 GGTGCTGCAAGCAACATCATCAAGAGATCCAAAGATCAAGAGATCAAGAGAT 1423  
DB 1381 GGTGCTGCAAGCAACATCATCAAGAGATCCAAAGATCAAGAGATCAAGAGAT 1423

RESULT 10  
AC36223  
ID AC36223 standard; DNA; 1431 BP.  
XX  
AC36223;  
XX  
AC36223;  
XX  
19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #17880.  
XX  
KM Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX drug design; gene.  
XX  
OS Klebsiella pneumoniae.  
XX  
PN K0200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002MO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX



PA (BLIT-) ELITRA PHARM INC.

XX Mang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JM;

PI Mail D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU32353.

XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 24093; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1431 BP; 285 A; 407 C; 480 G; 259 T; 0 U; 0 Other;

XX Query Match 70.2%; Score 1007; DB 8; Length 1431;

XX Best Local Similarity 81.7%; Pred. No. 1.8e-288;

XX Matches 1163; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 1 ATGAAAGTAAAGCTGCGGAGATTGAAAGTGCAGAGATGCTGTGGTGAATGATG 60  
DB 1 ATGAAAGTAAAGCTGCGGAGATTGAAAGTGCAGAGATGCTGTGGTGAATGATG 60  
QY 61 CTGAGCTGTTAAGGCGGAGCGCCCAAGTGTCTGCGCGAAGCGCGGTCGCCGTT 120  
DB 61 CTGAGCTGTTAAGGCGGAGCGCCCAAGTGTCTGCGCGAAGCGCGGTCGCCGTT 120  
QY 121 GTTAAAGTAAATCCATGAAAGAGTCCGGGCGGCGGCTTAAAGTGGCGATGATTC 180  
DB 121 GTTAAAGTAAATCCATGAAAGAGTCCGGGCGGCGGCTTAAAGTGGCGATGATTC 180  
QY 181 GCTTCTCTCGGTGCTAATGCAAGCTGTGGGTTGACGGGCTTTGACGATGCAAGCGCGC 240  
DB 181 GCTTCTCTCGGTGCTAATGCAAGCTGTGGGTTGACGGGCTTTGACGATGCAAGCGCGC 240  
QY 241 GCGGTGAGTAAATCTGTGGCGAGCTGCAAGCTCAAAATGCACTTGTCTGTGACGAGC 300  
DB 241 GCGGTGAGTAAATCTGTGGCGAGCTGCAAGCTCAAAATGCACTTGTCTGTGACGAGC 300  
QY 301 CATCGACATTAACAAATTAAGGGTACTTCCCGCAACAAAGCTGATCCGTCTGAT 360

DB 301 CACCCGACTATCACCAAGCTGCGGCTGTGTCCCGCAATCAACAGCTGAATCCCTCGAC 360  
QY 361 TTTGAAGAGGTTTCGAAGGTGTGATTCGACACCGCTGCAAGAGCGATTAATCAAGCG 420  
DB 361 TTTGAAGAGGCTTCTCCCGCTGCAATCCGACCCATGCAAGAGCGATTAATCAAGCG 420  
QY 421 CTGAGTTCGATTTGCGCGCTGTGTCTTTCGATCAAGCCAAAGGTGCGTCAAGCGTA 480  
DB 421 CTGAGTTCGATTTGCGCGCTGTGTCTTTCGATCAAGCCAAAGGTGCGTCAAGCGTA 480  
QY 481 CAGCAGATGATCAACTGCGCGCTAAAGCGGCTGTTCCGCTGATTTCAAAAAGT 540  
DB 481 CAGCAGATGATCAACTGCGCGCTAAAGCGGCTGTTCCGCTGATTTCAAAAAGT 540  
QY 541 ACCGATTTGAGGCTCAACCGCGCGCTAGCTGTAAACGCGCAATCTCGCAATTTGAA 600  
DB 541 ACCGATTTGAGGCTCAACCGCGCGCTAGCTGTAAACGCGCAATCTCGCAATTTGAA 600  
QY 601 GCTGTTGTCGTTAATGTAAGACCGAAGAAAGATTTGAGCGCGCATGAAATCTGATT 660  
DB 601 GCTGTTGTCGTTAATGTAAGACCGAAGAAAGATTTGAGCGCGCATGAAATCTGATT 660  
QY 661 GCGGATTAAGAACTCTGCGCTCTGTTAATGATCCGTTCCGAACAGGATATGCTGCTG 720  
DB 661 GCGGATTAAGAACTCTGCGCTCTGTTAATGATCCGTTCCGAACAGGATATGCTGCTG 720  
QY 721 CAACCGGGTAAAGCGCGCTGCAATGCAACCAAGCGCAGGAGTGTAGAGTTACC 780  
DB 721 CAACCGGGTAAAGCGCGCTGCAATGCAACCAAGCGCAGGAGTGTAGAGTTACC 780  
QY 781 GGTGCGGCGCAACGCTGATTTGCGCTCTGCGCGCAACGCTGCAAGCGGTTAATTGCTG 840  
DB 781 GGTGCGGCGCAACGCTGATTTGCGCTCTGCGCGCAACGCTGCAAGCGGTTAATTGCTG 840  
QY 841 GAAAGAGCTGCTTCTTTGCAATGCGGCGGCTGCGGTGTGTCTGCGCAATCTGGAACC 900  
DB 841 GAAAGAGCTGCTTCTTTGCAATGCGGCGGCTGCGGTGTGTCTGCGCAATCTGGAACC 900  
QY 901 TTCACGGTTTCGCGAATCGAGCTGGAATAATGCTGATGCTGCAAGCTGCAATGAGGCTTT 960  
DB 901 TTCACGGTTTCGCGAATCGAGCTGGAATAATGCTGATGCTGCAAGCTGCAATGAGGCTTT 960  
QY 961 GCGGTGATGACCGAAGAGAACTGAAGCTGCGGTGACGCGCAGCGCTGAACCTGTGAA 1020  
DB 961 GCGGTGATGACCGAAGAGAACTGAAGCTGCGGTGACGCGCAGCGCTGAACCTGTGAA 1020  
QY 1021 AAAGTGTATGACCAACGCTGCTTTGACATCTGCAAGCGCGGCAAGCTCTTATCTG 1080  
DB 1021 AAAGTGTATGACCAACGCTGCTTTGACATCTGCAAGCGCGGCAAGCTCTTATCTG 1080  
QY 1081 GCAAAATGCCGCAAGCTGGGTGAACGCTTGAATTTGTCGCTCAACAGCGATGCTCCACC 1140  
DB 1081 GCAAAATGCCGCAAGCTGGGTGAACGCTTGAATTTGTCGCTCAACAGCGATGCTCCACC 1140  
QY 1141 AAACGCTGTAAGAGCGAAATCCGCGCGGTGAACCACTCGAACAAGCTGATGATTTGCTG 1200  
DB 1141 AAACGCTGTAAGAGCGAAATCCGCGCGGTGAACCACTCGAACAAGCTGATGATTTGCTG 1200  
QY 1201 GCGGCACTGGAAGCGGTGCACTGCGGTGATGCTGTTGAAGAAGACACGCGCAGCGCTTG 1260  
DB 1201 GCGGCACTGGAAGCGGTGCACTGCGGTGATGCTGTTGAAGAAGACACGCGCAGCGCTTG 1260  
QY 1261 ATGCGCGGAGATCTTGGCAAGATCTGCTGCTGTAAGAAAGCGCGCACTTAACAGAGAATTT 1320  
DB 1261 ATGCGCGGAGATCTTGGCAAGATCTGCTGCTGTAAGAAAGCGCGCACTTAACAGAGAATTT 1320  
QY 1321 GCGGAGTAAAGAAAGTCTGCGCAACGCTGCGGAAGTGTGTGCTCAACTTTGAAGAC 1380  
DB 1321 GCGGAGTAAAGAAAGTCTGCGCAACGCTGCGGAAGTGTGTGCTCAACTTTGAAGAC 1380  
QY 1381 GGTGCTGACGACCAATCATCAAGAAATGCCAAGATTA 1423







Db 481 CAGCCGATGATCCAACTGCGACGTAAGCCAAATGTCCTGGTTTGTATTCGAAAGGC 540  
 Qy 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTTTAAAGCGCAATCTCTCGAAATTTGAA 600  
 Db 541 AGTGAATTTGAGCGCTACCGCTGCGTGCACATTTGTTGACGCGCAATTTGTCTGAATTTGAA 600  
 Qy 601 GCTGTTTTCGTTAAATTTAAGACCGAAGAAAGATTTTGAAGCGCGCATGAAATCTGAT 660  
 Db 601 GCAATGTTAGGCGCTGTAAATAATGAAGAAAGTTGTTAAACCGGTAATGCAACTGCTG 660  
 Qy 661 GCGCATACGAATCTCGGCTCTGTATGTAACCGCTTCCGAACAAGGATATGCTGCTG 720  
 Db 661 GCAGATTTTGAATCTTATGATCTGCTGTAACCGCTTCTGAACAAGGATGACCTTGTG 720  
 Qy 721 CAACCGGATTAAGCGCGCTGCAATATGCAACCGCAAGCGAAGATGTAATGATTAAC 780  
 Db 721 CAACCTGGATTAAGCGCGCTGCAATATGCAACCGCAAGCGAAGATGTAATGATTAAC 780  
 Qy 781 GGTGCGGCGGCAACGATGATTTGGGCTCTGCGCGCAACGCTGCGAGCGGGTAAATTCGCTG 840  
 Db 781 GGTGCTGCGCAATACGATTAATTTGGGCTCTGCGCGCAACGCTGCGAGGATTAACGCTG 840  
 Qy 841 GAAGAAGCTGCTCTTCTTTCGAATGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTG 900  
 Db 841 GAAGAAGCTGCTCTTCTTTCGAATGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTG 900  
 Qy 901 TCCAGGTTTTCGCGGATGAGCTGGAATGCTGTACGTGACGTGCAATACAGGCTTT 960  
 Db 901 TCCAGGTTTTCGCGGATGAGCTGGAATGCTGTACGTGACGTGCAATACAGGCTTT 960  
 Qy 961 GCGCTGATGACCGAAGGAATGCAATGCTGCGCTGCGGCGGCTGCGGCTGCGGCTG 1020  
 Db 961 GCGCTGATGACCGAAGGAATGCAATGCTGCGCTGCGGCGGCTGCGGCTGCGGCTG 1020  
 Qy 1021 AAGGTGATGATGACCAACGCTGCTTGTACATCTGTGACGCGCGGCAAGCTCTTATCTG 1080  
 Db 1021 AAGGTGATGATGACCAACGCTGCTTGTACATCTGTGACGCGCGGCAAGCTCTTATCTG 1080  
 Qy 1081 GCAAAATCCCGCAAGCTGCGTACCGCTTGTATGTTGCTGCAACGAGCTCTCAAC 1140  
 Db 1081 GCAAAATCCCGCAAGCTGCGTACCGCTTGTATGTTGCTGCAACGAGCTCTCAAC 1140  
 Qy 1141 AAGCGCTGAAAGGAGATTCGCGCGGCTAAACCACTGCAACAGCGTAAATGATGCTG 1200  
 Db 1141 AAGCGCTGAAAGGAGATTCGCGCGGCTAAACCACTGCAACAGCGTAAATGATGCTG 1200  
 Qy 1201 GCGCACTGAAAGGAGATTCGCGCGGCTAAACCACTGCAACAGCGTAAATGATGCTG 1260  
 Db 1201 GCGCACTGAAAGGAGATTCGCGCGGCTAAACCACTGCAACAGCGTAAATGATGCTG 1260  
 Qy 1261 ATGCGCGGATCTTTCGCAATCTGCTGCTGTAAGAGCGCGCACTTAAACCAAGAGAT 1320  
 Db 1261 ATGCGCGGATCTTTCGCAATCTGCTGCTGTAAGAGCGCGCACTTAAACCAAGAGAT 1320  
 Qy 1321 GCGCGAGATTAAGAGATTCGCGCGGCTAAACCACTGCAACAGCGTAAATGATGCTG 1380  
 Db 1321 GCGCGAGATTAAGAGATTCGCGCGGCTAAACCACTGCAACAGCGTAAATGATGCTG 1380  
 Qy 1381 GGTGCTGCAAGCAATCATCAAGAGATTCGAACG 1420  
 Db 1381 GGTGCTGCAAGCAATCATCAAGAGATTCGAACG 1420

RESULT 13  
 ACH96312/c  
 ID ACH96312 standard, DNA; 1323 BP.

XX ACH96312;  
 XX 29-JUL-2004 (first entry)  
 XX Klebsiella pneumoniae polynucleotide seqid 2107.  
 XX

KW Recombinant expression vector; transcription regulatory element;  
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.  
 OS Klebsiella pneumoniae.  
 PN US6610836-B1.  
 PD 26-AUG-2003.  
 PF 27-JAN-2000; 2000US-00489039.  
 PR 29-JAN-1999; 99US-0117747P.  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 PI Breton GL, Osborne M;  
 DR WPI: 2003-895346/82.  
 DR P-PSDB; ABO62761.  
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for  
 PS preparing a vaccine composition against Klebsiella pneumoniae.  
 CC Disclosure; SEQ ID NO 2107; 932pp; English.  
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella  
 CC pneumoniae polypeptide. Also described are: a recombinant expression  
 CC vector comprising the nucleic acid, operably linked to a transcription  
 CC regulatory element; and a cell comprising the recombinant expression  
 CC vector. The nucleic acid is useful for preparing a vaccine composition  
 CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella  
 CC pneumoniae polypeptide of the invention  
 XX  
 SQ Sequence 1323 BP; 244 A; 450 C; 376 G; 253 T; 0 U; 0 Other;  
 Query Match 60.7%; Score 870.6; DB 11; Length 1323;  
 Best Local Similarity 81.8%; Pred. No. 6.6e-248;  
 Matches 1005; Conservative 0; Mismatches 224; Indels 0; Gaps 0;  
 Qy 1 ATGAAAGTAAACGCTGCAAGATTTGAACGTGCAAGATGATGCTGATGATGATG 60  
 Db 1230 ATGAAAGTAAACGCTGCAAGATTTGAACGTGCAAGATGATGCTGATGATGATG 1171  
 Qy 61 CTGATGCTTACCTGATGCGGCCCAACGATGCTGCTGCGCGGAACCGCGTCCGCTG 120  
 Db 1170 CTGATGCTTACCTGATGCGGCCCAACGATGCTGCTGCGCGGAACCGCGTCCGCTG 1111  
 Qy 121 GTTAAAGTAAATCAATCGAAGAAAGTCCGCGCGCGCGCTAAACGTGCGATGATATC 180  
 Db 1110 GTTAAAGTAAATCAATCGAAGAAAGTCCGCGCGCGCGCTAAACGTGCGATGATATC 1051  
 Qy 181 GCTTCTCTCGGCTATGATGACCGCTGCTGCGGCTTGAAGCGGCAATGACATGACGCGC 240  
 Db 1050 GCTTCTCTCGGCTATGATGACCGCTGCTGCGGCTTGAAGCGGCAATGACATGACGCGC 991  
 Qy 241 GCGCTGATTAATCTCTGCGCGGCTGACGATGCAATGCAATGCAATGCAATGCAATG 300  
 Db 990 GCGCTGATTAATCTCTGCGCGGCTGACGATGCAATGCAATGCAATGCAATGCAATG 931  
 Qy 301 CATCGACATTAACCAATTAAGGATGATTTCCGCGAACAACGATGATCCGCTGAT 360  
 Db 930 CATCGACATTAACCAATTAAGGATGATTTCCGCGAACAACGATGATCCGCTGAT 871  
 Qy 361 TTGAAAGAGTTTGAAGGATGATTCGCAACCGCTGACGAGCGGATTAATCAAGCG 420  
 Db 870 TTGAAAGAGTTTGAAGGATGATTCGCAACCGCTGACGAGCGGATTAATCAAGCG 811  
 Qy 421 CTGAGTTCAATTCGCGCGCTGCTGCTTCTGATTAAGCGCAAGATGCTGCGAAGCTA 480  
 Db 810 CTGAGTTCAATTCGCGCGCTGCTGCTTCTGATTAAGCGCAAGATGCTGCGAAGCTA 751  
 Qy 481 CAGCAATGATTCGAATGCGCGCTAAAGCGGATGCTCCGCTGATGATGCAAAAGT 540

```
Dh 750 CAGACATGATCCGACTCGACGTGAAGCGGCGTGCCTGATCATCCGAAAGGC 691
Qy 541 ACCGATTTTGAGCGCTACCGGCGGCTACGCTTTAAACCGAAATCTCTGGAAATTTGAA 600
Dh 690 ACCGATTTTGAGCGCTACCGGCGGCTACGCTTTAAACCGAAATCTCTGGAAATTTGAA 631
Qy 601 GCTGTGTGCTGTAATATGATGACCGAAGAGATTTGTAAGCGCGGATGAACTGATTT 660
Dh 630 GCGGTGAGTGGCAAGTGCAGAGATGAAGCGAGATCGTTGAGCGCGGATGAACTGATTT 571
Qy 661 GCCGATTAAGAACTCTGGGCTCTGTGATGACCGGCTTCCGAAACGAGTATGTCCTGCTG 720
Dh 570 GCCGATTAAGAACTCTGGGCTCTGTGATGACCGGCTTCCGAAACGAGTATGTCCTGCTG 511
Qy 721 CAACCGGCTAAAGCGCGCTGATATGCAACCGAAGCGGAGAGTATGATGAGTTACC 780
Dh 510 CAGCGCGGAGCTCCGCGCTGATATGCAACCGAAGCGGAGAGTATGATGAGTTACC 451
Qy 781 GGTGCGGCGGACACGCGTATTTGGCTCTGCGGCAACGCTGAGCGGATTTGCTG 840
Dh 450 GGGCGCGGCGACACGCGTATTTGGCTCTGCGGCAACGCTGAGCGGATTTGCTG 391
Qy 841 GAAGAAGCTGCTCTTTTCCGAAATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Dh 390 GAAGAAGCTGCTCTTTTCCGAAATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 331
Qy 901 TCCAGGCTTTCGCGGATGAGCTGGAATAATGCTGATGAGTATGAGTATGAGTATGAGTAT 960
Dh 330 TCCAGGCTTTCGCGGATGAGCTGGAATAATGCTGATGAGTATGAGTATGAGTATGAGTAT 271
Qy 961 GGCCTGATGACGGAAGAGAACTGAAGCTGCGCTGAGCGGAGCGCTGAAACGTTGTTGAA 1020
Dh 270 GGCCTGATGACGGAAGAGAACTGAAGCTGCGCTGAGCGGAGCGCTGAAACGTTGTTGAA 211
Qy 1021 AAATGCTGATGACCAACGCTGCTTTGACATCTCTGACGCGCGGAGCTCTCTTATCTG 1080
Dh 210 AAATGCTGATGACCAACGCTGCTTTGACATCTCTGACGCGCGGAGCTCTCTTATCTG 151
Qy 1081 GCAATATGCCGCAAGCTGAGTGAACGCTTGAATTTGTCCTGCAACAGAGATGCTCCAGC 1140
Dh 150 GCAATATGCCGCAAGCTGAGTGAACGCTTGAATTTGTCCTGCAACAGAGATGCTCCAGC 91
Qy 1141 AAACGCTGAAAGGAGATTTCCGCGCTGAAACCACTCGAAGCGGTATGATTTGCTG 1200
Dh 90 AAACGCTGAAAGGAGATTTCCGCGCTGAAACCACTCGAAGCGGTATGATTTGCTG 31
Qy 1201 GCGCAGCTGAAAGGCTGATGAGTGTAGT 1229
Dh 30 GCGCAGCTGAAAGGCTGATGAGTGTAGT 2
```

## RESULT 14

ACF67367\_45  
Continuation (46 of 57) of ACF67367 from base 4500001 (Photobacterium luminescens nucleotid  
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

Fragment Name	Begin	End
ACF67367_00	1	110000
WP ACF67367_01	100001	210000
WP ACF67367_02	200001	310000
WP ACF67367_03	300001	410000
WP ACF67367_04	400001	510000
WP ACF67367_05	500001	610000
WP ACF67367_06	600001	710000
WP ACF67367_07	700001	810000
WP ACF67367_08	800001	910000
WP ACF67367_09	900001	1010000
WP ACF67367_10	1000001	1110000
WP ACF67367_11	1100001	1210000
WP ACF67367_12	1200001	1310000
WP ACF67367_13	1300001	1410000
WP ACF67367_14	1400001	1510000
WP ACF67367_15	1500001	1610000
WP ACF67367_16	1600001	1710000

WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894

Query Match 54.7%; Score 784.8; DB 10; Length 110000;

Best Local Similarity 71.8%; Pred. No. 1,8e-221; Mismatches 402; Indels 0; Gaps 0;

Matches 1026; Conservative 0;

Qy	1	ATGAAGTAAAGCTGCGCAGAGTTGAAACGTGACGAGATGAGTGGTGGATGATGATG 60
Dh	38374	ATGAAGTAAAGCTGCGCAGAGTTGAAACGTGACGAGATGAGTGGTGGATGATGATG 38433
Qy	61	CTGATGCTTACTGTATGACGCGCCACACAGTCTATCTCGCGGAGCGCGGTCGCGT 120
Dh	38434	TTAGATGCTACTGATGATGAGCGGACGACGCGGATTTCCGCGGAGCGCGGTCGCTGA 38493
Qy	121	GTTAAAGTAAATCCATCGAAGAACTCCGCGGCGCGGCTTAAAGTGGGATGAATATC 180
Dh	38494	GTTAAAGTAAATCCATCGAAGAACTCCGCGGCGCGGCTTAAAGTGGGATGAATATC 38553
Qy	181	GCTTCTCTGCGTATGATGACGCGCTGATCGGTTGACGCGGATTAAGATGACGCGCG 240
Dh	38554	GCTTCTCTGCGTATGATGACGCGCTGATCGGTTGACGCGGATTAAGATGACGCGCG 38613
Qy	241	GCGCTGATTAATCTTGGCGGACGTCAAGTCAATGCGATTCGTTCTGTAACGACG 300
Dh	38614	GCGCTGATTAATCTTGGCGGACGTCAAGTCAAGTCAATGCGATTCGTTCTGTAACGACG 38673
Qy	301	CATTCGACCATTTCCAAATTAACGGGATCTTCCCGCAACCAAGCTGATCCGCTGAT 360
Dh	38674	CATTCGACCATTTCCAAATTAACGGGATCTTCCCGCAACCAAGCTGATCCGCTGAT 38733
Qy	361	TTTGAAGAGGTTTCGAAGGTTGATCCGACGCGCTGACGAGCGGATTAATCAGGCG 420
Dh	38734	TTTGAAGAGGTTTCGAAGGTTGATCCGACGCGCTGACGAGCGGATTAATCAGGCG 38793

[illegible]

Db 85843 AGTCTGATCAGCCACACCTGCATTTGCCACGACGAGCCGACAGAACTGTTTATGTATAG 8578

Gy 781 GGTGCGGCGACACAGGATTTGGCGTCTGTGCGGCAACGCTGGACGCGGTATTTGGCTG 840

Db 85783 GGTGCTGCGCATACGATTATGTGTGTAGCAACGCGCATTCGCGCTGGAAGCCACTG 85724  
QY 841 GAAAGAACCTGCTCTTTGCCAATGCGCGCGCTGCGTGTGCGGCAAACTGGGAACC 900  
Db 85723 AATGAAGCTGTCTTCTGCGCAATGCTGCGGCGGTGTGTGTGGCAAACTGGGAACG 85664  
QY 901 TCCACGGTTTCGCCGATCGAGCTGAAATGCTGTAGCTGACGTGACGTGACATACAGGCTTT 960  
Db 85663 TCGACCGTTTCGCCGATTTGAACTGGAATGCTATCCGTGCGCGCGTGAACAGGTTTC 85604  
QY 961 GCGCTGATGACCGAAGAGAACTGAAGTGGCGCTAGCGGCAAGCGCTMAAAGTGTGAA 1020  
Db 85603 GGTGTGATGACCGAATTCAGCTTAAACAGGCTGTGTGATGCTCGTCAACGTGTGAA 85544  
QY 1021 AAAGTGTGATGACCAACGCGTCTTTGACATCTGCAACGCGCGGCAAGTCTCTTATCTG 1080  
Db 85543 AGGTTGTGATGACTAATGCTGCTTTGATATCTGCAATGCAAGGTCAATGCTCTTATCTG 85484  
QY 1081 GCAATGCCCCGCAAGCTGGTGAACGCTTGAATGTGCGGTCAACAGCGATGCTTCACC 1140  
Db 85483 GAGAAATGCCGTAGACTGGGTGATCGGTAAATGTGTGTGTAACAGTATGATCAACT 85424  
QY 1141 AAACGCTGAAAGGGGATTCGCCGCCGCTAAACCACTCGAACAGCGTATGATGTGCTG 1200  
Db 85423 AAACGCTGAAAGGTGAGAGTGTCTCAGTTAATCAGTAAACAGAAATGATGTGTTA 85364  
QY 1201 GCGGCACTGGAAGGGGTGCACTGGGTAGTGTGTTTGAAGAGGACAGCGCGCAGCGCTTG 1260  
Db 85363 TCTGCTGCGCGCGGTGATTTGGGTGTTCATTGTGAAGAGATACCAACAGCGAATG 85304  
QY 1261 ATCCCGCGGATCTTGCCAGATCTGTGTTGAAGCGCGCATATMAACGAGAGATT 1320  
Db 85303 ATTGCGGATGTTTTGCCGAGTGTATGTGTAAAGGGGAGATTATMAACCGGAGAAATT 85244  
QY 1321 GCCGGAGTAAAGAGTCTGGGCCAAGGTGGCGAAGTGTGCTCACTTTGAAGAC 1380  
Db 85243 GCTGTAGTGAAGAGTTTGGGCTGCGGCGGAGACGTCAAAAGTTGAACTTTGAAAGAT 85184  
QY 1381 GGTGCTGACGACCAATCATCATMAAGATCCACAGATTAATAA 1428  
Db 85183 GGTATTTGACACCAATATATCATCAAGGCGATTMAAATCAGTAAATA 85136

Search completed: March 18, 2006, 19:07:31  
Job time : 945 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2006, 05:59:13 ; Search time 188 seconds

(without alignments)  
1114.807 Million cell updates/sec

Title: US-10-520-820-13

Sequence: 1 MKXTLPFRFAGAGVWVGVDM.....FEDGCTNIIKKIQDDKKG 477

Scoring table: BIOSUM62

Gap 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

1: A\_GeneSeq\_21:\*  
2: geneeqp19808:\*  
3: geneeqp19908:\*  
4: geneeqp20008:\*  
5: geneeqp20018:\*  
6: geneeqp20028:\*  
7: geneeqp20038:\*  
8: geneeqp20048:\*  
9: geneeqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2393	100.0	477	3	ABU15968
2	2393	100.0	477	4	AAU34708
3	2393	100.0	477	4	AAU34708
4	2393	100.0	477	6	ABU14690
5	2393	100.0	477	6	ABU14690
6	2294	95.9	476	6	ABU28278
7	2265	94.7	477	6	ABU28166
8	2265	94.7	477	6	ABU47554
9	2227	93.1	477	6	ABU32353
10	2227	93.1	477	6	ABU32353
11	2196	91.8	463	6	ABU45244
12	2076	86.8	476	6	ABU50241
13	2007	83.9	474	6	ABU40796
14	2007	83.9	491	7	ADP06349
15	1950	81.5	475	6	ABU67735
16	1712	71.1	476	6	ABU36009
17	1701	71.1	476	6	ABU39200
18	1357.5	56.7	474	6	ABU41388
19	1351.5	56.5	525	7	ABU084238
20	1350.5	56.4	474	4	AAU36486
21	1350.5	56.4	474	6	ABU38845
22	1338.5	55.9	469	6	ABU39925
23	863	36.1	461	6	ABU26598
24	825	34.5	334	6	ABU21555

25	809.5	33.8	461	6	ABU30896	ABU30896 Protein e
26	745.5	31.2	323	5	AAU72936	AAU72936 Neisseria
27	745.5	31.2	323	6	ABU37928	ABU37928 Protein e
28	744.5	31.1	320	6	ABP78362	ABP78362 N. gonorr.
29	744.5	31.1	320	6	ABU37300	ABU37300 Protein e
30	734.5	30.7	424	2	AAW20763	AAW20763 H. pylori
31	732.5	30.6	313	6	ABU22903	ABU22903 Protein e
32	691.5	28.9	316	6	ABU19695	ABU19695 Protein e
33	679.5	28.4	328	6	ABU22239	ABU22239 Protein e
34	610.5	25.5	428	3	ABU07577	ABU07577 Protein e
35	564.5	23.6	315	8	ADN17384	ADN17384 Bacterial
36	490.5	20.5	312	5	ABU50893	ABU50893 Helicobac
37	420	17.6	315	8	ADS21181	ADS21181 Bacterial
38	362	15.1	161	6	ABU22547	ABU22547 Protein e
39	354.5	14.8	187	5	ABU50751	ABU50751 Helicobac
40	350	14.6	161	6	ABU20231	ABU20231 Protein e
41	344	14.4	178	6	ABU21806	ABU21806 Protein e
42	341.5	14.3	168	6	ABU37318	ABU37318 Protein e
43	341.5	14.3	169	6	ABP77009	ABP77009 N. gonorr
44	340.5	14.2	168	6	ABU37765	ABU37765 Protein e
45	332	13.9	162	6	ABU23023	ABU23023 Protein e

## ALIGNMENTS

RESULT 1	AAU15968	standard; protein; 477 AA.
AC	AAU15968;	
DT	05-OCT-2000	(first entry)
DE	E. coli proliferation associated protein sequence SEQ ID NO:325.	
XX	Bacterichia coli; B. coli; proliferation; inhibition; screening;	
KW	antimicrobial; bacterial growth; antisense therapy; antibacterial.	
OS	Bacterichia coli.	
XX	WO200044906-A2.	
PN	03-AUG-2000.	
PD	27-JAN-2000; 2000MO-US002200.	
XX	27-JAN-1999; 99US-0117405P.	
PR	(ELIT-) ELITRA PHARM INC.	
PA	Zyckind J, Ohlsen KL, Trawick J, Forsyth RA, Froelich JM;	
PI	Carr GT, Yamamoto RT, Xu HH;	
XX	WPI; 2000-514822/46.	
DR	N-PSDB; AAU55973.	
XX	Novel polynucleotides and polypeptides associated with microorganism	
PT	proliferation, used to identify inhibitors of bacterial growth and	
PT	proliferation, for use in antisense therapy.	
XX	Claim 11; Page 242-243; 316pp; English.	
PS	AAU5809 to AAU5889 and AAU6058 to AAU6138 represent nucleotide	
XX	sequences derived from Bacterichia coli which inhibit E. coli	
CC	proliferation. AAU5890 to AAU6055 and AAU15886 to AAU16040 represent	
CC	nucleotide and protein sequences associated with E. coli proliferation.	
CC	AAU6056 and AAU6057 represent primers used for sequencing E. coli	
CC	proliferation inhibiting nucleotide inserts in an example from the	
CC	present invention. Methods from the present invention can be used to	
CC	identify a proliferation-regulated gene in a microorganism, by contacting	
CC	a microorganism with a proliferation-regulated gene activity inhibitory	
CC	nucleic acid identified in another organism, and determining if	



inhibition occurs in the second microorganism. The nucleic acid sequences identified as being required for bacterial growth and proliferation, can be used for antisense therapy for killing bacteria

Sequence 477 AA:

Query Match 100.0%; Score 2393; DB 3; Length 477;  
Best Local Similarity 100.0%; Pred. No. 6.9e-207;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKVTLPFERBAGVWVGVMDLDRVYGTPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 60
DB 1 MKVTLPFERBAGVWVGVMDLDRVYGTPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 60
QY 61 ASLGANARLVGLTGIDDAARALSKSLADVANKCPFVSPHTPTTKLRVLSRNOQLRLD 120
DB 61 ASLGANARLVGLTGIDDAARALSKSLADVANKCPFVSPHTPTTKLRVLSRNOQLRLD 120
QY 121 FEEGFEVDPPQLHBRINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPVLIDPKG 180
DB 121 FEEGFEVDPPQLHBRINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPVLIDPKG 180
QY 181 TDFERYRGATLTLPNTLSEFEAVVCKTEBEIYERGMKLIADYELSLALVTRSQGMSLL 240
DB 181 TDFERYRGATLTLPNTLSEFEAVVCKTEBEIYERGMKLIADYELSLALVTRSQGMSLL 240
QY 241 QPGKAPLHMPTOAQEVVDVDTAGDVTYIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
DB 241 QPGKAPLHMPTOAQEVVDVDTAGDVTYIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
QY 301 STVSPLEENAVRGRADTGFVMTBEELKLAVAAARKGKRVMTNGVFDILHAGHVSYL 360
DB 301 STVSPLEENAVRGRADTGFVMTBEELKLAVAAARKGKRVMTNGVFDILHAGHVSYL 360
QY 361 ANARKLGDRLIVAVNSDASTKRLKGSRPVNPLEORMIVLGALRAVDMVVSFEEDTQRL 420
DB 361 ANARKLGDRLIVAVNSDASTKRLKGSRPVNPLEORMIVLGALRAVDMVVSFEEDTQRL 420
QY 421 IAGILPDLVKGDDYKPEEIASKEVWANGGEVLVNFEDGCSITNIIKKIQODKKG 477
DB 421 IAGILPDLVKGDDYKPEEIASKEVWANGGEVLVNFEDGCSITNIIKKIQODKKG 477

RESULT 2
ID AU34708 standard; protein; 477 AA.
AC AU34708;
XX 14-FEB-2002 (first entry)
DT 14-FEB-2002 (first entry)
XX E. coli cellular proliferation protein #289.
DE E. coli cellular proliferation protein #289.
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX Escherichia coli.
OS Escherichia coli.
XX MO200170955-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009180.
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207727P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.

```

XX Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ,  
PI Yamamoto RT, Xu HH,  
XX WPI, 2001-611495/70.  
DR N-PSDB; AAS52567.  
XX New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.  
XX

PS Example 3; SEQ ID NO 10301; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*  
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 477 AA;

Query Match 100.0%; Score 2393; DB 4; Length 477;  
Best Local Similarity 100.0%; Pred. No. 6.9e-207;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKVTLPFERBAGVWVGVMDLDRVYGTPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 60
DB 1 MKVTLPFERBAGVWVGVMDLDRVYGTPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 60
QY 61 ASLGANARLVGLTGIDDAARALSKSLADVANKCPFVSPHTPTTKLRVLSRNOQLRLD 120
DB 61 ASLGANARLVGLTGIDDAARALSKSLADVANKCPFVSPHTPTTKLRVLSRNOQLRLD 120
QY 121 FEEGFEVDPPQLHBRINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPVLIDPKG 180
DB 121 FEEGFEVDPPQLHBRINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPVLIDPKG 180
QY 181 TDFERYRGATLTLPNTLSEFEAVVCKTEBEIYERGMKLIADYELSLALVTRSQGMSLL 240
DB 181 TDFERYRGATLTLPNTLSEFEAVVCKTEBEIYERGMKLIADYELSLALVTRSQGMSLL 240
QY 241 QPGKAPLHMPTOAQEVVDVDTAGDVTYIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
DB 241 QPGKAPLHMPTOAQEVVDVDTAGDVTYIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
QY 301 STVSPLEENAVRGRADTGFVMTBEELKLAVAAARKGKRVMTNGVFDILHAGHVSYL 360
DB 301 STVSPLEENAVRGRADTGFVMTBEELKLAVAAARKGKRVMTNGVFDILHAGHVSYL 360
QY 361 ANARKLGDRLIVAVNSDASTKRLKGSRPVNPLEORMIVLGALRAVDMVVSFEEDTQRL 420
DB 361 ANARKLGDRLIVAVNSDASTKRLKGSRPVNPLEORMIVLGALRAVDMVVSFEEDTQRL 420
QY 421 IAGILPDLVKGDDYKPEEIASKEVWANGGEVLVNFEDGCSITNIIKKIQODKKG 477
DB 421 IAGILPDLVKGDDYKPEEIASKEVWANGGEVLVNFEDGCSITNIIKKIQODKKG 477

```

RESULT 3  
ID AAG98421 standard; protein; 477 AA.

XX AAG98421;  
 AC  
 XX 21-SEP-2001 (first entry)  
 DT  
 XX Escherichia coli protein sequence SEQ ID NO:469.  
 DE  
 XX Escherichia coli; identification; proliferation; microorganism;  
 KW antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;  
 KW bacterial growth inhibition.  
 XX Escherichia coli.  
 OS  
 XX NC0200148209-A2.  
 PN  
 XX 05-JUL-2001.  
 PD  
 XX 19-DEC-2000; 2000MO-US034419.  
 PF  
 XX 23-DEC-1999; 99US-0173005P.  
 PR  
 XX (ELIT-) ELITRA PHARM INC.  
 PA  
 XX Forsyth RA, Ohlsen KL, Zyskind JW;  
 PI WPI, 2001-457376/49.  
 XX N-PSDB; AAH81477.  
 DR  
 XX Novel nucleic acids encoding proteins required for Escherichia coli  
 PT proliferation, useful for screening for antimicrobial agents.  
 XX  
 XX Example 4, Page 585-586; 596pp; English.  
 PS  
 XX The present invention describes a purified or isolated nucleic acid  
 CC sequence (1) consisting essentially of one of the 93 nucleotide sequences  
 CC given in AAH81202 to AAH81294, where expression of the nucleic acid in a  
 CC microorganism is capable of inhibiting proliferation of a microorganism.  
 CC (1) have antibacterial and antibiotic activities, and can be used in gene  
 CC therapy. Expression of (1) in a microorganism inhibits proliferation of  
 CC the microorganism, and the manufactured antibiotic is useful for reducing  
 CC the activity or level of a gene product required for proliferation of a  
 CC microorganism in a subject, specifically humans. The nucleic acids that  
 CC inhibit bacterial growth or proliferation can be used as antisense  
 CC therapeutics for killing bacteria. In addition to therapeutic  
 CC applications, the nucleic acid sequences complementary to sequences  
 CC required for proliferation can be used as diagnostic tools. For example,  
 CC nucleic acid probes complementary to proliferation-regulated sequences  
 CC that are specific for particular species of microorganisms can be used as  
 CC probes to identify particular microorganism species in clinical  
 CC specimens. AAH81295 to AAH81487 encode the Escherichia coli proteins  
 CC given in AAG98239 to AAG98431, and AAH81488 to AAH81491 represent  
 CC oligonucleotides, which are used in the exemplification of the present  
 CC invention  
 CC  
 CC Sequence 477 AA;  
 SQ  
 Query Match 100.0%; Score 2393; DB 4; Length 477;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-207;  
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 181 TDPERYRGATLTPNLSPEAVVGVCKTEBEIVRGKMLADVELSLVLTRESGMSLL 240  
 QY 241 QPKAPLHMPPTQAEVYDVYAGDVIYGVLAATLAAGNSIEBAQCFPANAAGVVGGLGT 300  
 DB 241 QPKAPLHMPPTQAEVYDVYAGDVIYGVLAATLAAGNSIEBAQCFPANAAGVVGGLGT 300  
 QY 301 STVSPIELBNVAVGRADTGVGWTTEBEELKLAVAARRGKVMYMTGVPDIIAAGVSYL 360  
 DB 301 STVSPIELBNVAVGRADTGVGWTTEBEELKLAVAARRGKVMYMTGVPDIIAAGVSYL 360  
 QY 361 ANARKGDRLIYVNSDASTKRLKGSRPVNPLEORHIVGALBAVDWVVSFEEDTPORL 420  
 DB 361 ANARKGDRLIYVNSDASTKRLKGSRPVNPLEORHIVGALBAVDWVVSFEEDTPORL 420  
 QY 421 IAGILPDLAVKGGDYKPEBIAGSKVWANGGEVLVNFEDGCGSTNIKKIQDDKKG 477  
 DB 421 IAGILPDLAVKGGDYKPEBIAGSKVWANGGEVLVNFEDGCGSTNIKKIQDDKKG 477

RESULT 4  
 ABU14690  
 ID ABU14690 standard; protein; 477 AA.  
 XX  
 AC ABU14690;  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by prokaryotic essential gene #217.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Escherichia coli.  
 XX  
 PN NC0200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002MO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI, 2003-029926/02.  
 DR N-PSDB; ACA18560.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids, required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 25; SEQ ID NO 42614; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 623 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)

identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 477 AA:

Query Match 100.0%; Score 2393; DB 6; Length 477;  
Best Local Similarity 100.0%; Pred. No. 6.9e-207; Mismatches 0; Indels 0; Gaps 0;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKVTLPEPERAGVWVGVDMVMDRMYGPTSRISPEAPVVPVVKVNTIERRPGGANVAMNI 60  
1 MKVTLPEPERAGVWVGVDMVMDRMYGPTSRISPEAPVVPVVKVNTIERRPGGANVAMNI 60  
61 ASIGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRLVSRNOQLRLD 120  
61 ASIGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRLVSRNOQLRLD 120  
121 FEEGFEGVDPQPLHERINQALSSIGALVLSDYAKGALASVQOMIOLARKAGVPLIDPKG 180  
121 FEEGFEGVDPQPLHERINQALSSIGALVLSDYAKGALASVQOMIOLARKAGVPLIDPKG 180  
181 TDFERRYGATLTLPNTSEFAVVGKCKTEBEIYERGMKLIADVELSALVTRSEOGMSLL 240  
181 TDFERRYGATLTLPNTSEFAVVGKCKTEBEIYERGMKLIADVELSALVTRSEOGMSLL 240  
241 QPGKAPLHPMPTQAOBYVDVAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
241 QPGKAPLHPMPTQAOBYVDVAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
301 STVSPTELENAVGRADTGFVMTBEELKLAVAARRGKGVMTNGVFDIILHAGHVSYL 360  
301 STVSPTELENAVGRADTGFVMTBEELKLAVAARRGKGVMTNGVFDIILHAGHVSYL 360  
361 ANARKGDRLLIVAVNSDASTKRLKGSRPVNPLEORMIVGALBAVDVVSFEEDTPORL 420  
361 ANARKGDRLLIVAVNSDASTKRLKGSRPVNPLEORMIVGALBAVDVVSFEEDTPORL 420  
421 IAGILPDLVLVGGDYKPEEIASKEVWANGGEVLVNFEDGCSYTNIIKKIQODKKG 477  
421 IAGILPDLVLVGGDYKPEEIASKEVWANGGEVLVNFEDGCSYTNIIKKIQODKKG 477

RESULT 5  
ADH88979 standard; protein; 477 AA.

ADH88979;

22-APR-2004 (first entry)

Becherichia coli K1 S26 mutant, SEQ ID 13.

Pathogenic microorganism; K1 S26; mutant; pathogenicity; virulence; bacterial infection; extra-intestinal infection; enterobacterium.

Becherichia coli.

OS Synthetic.

MO2004005535-A2.  
15-JAN-2004.  
09-JUL-2003; 2003MO-EP008209.  
09-JUL-2002; 2002EP-0008636.  
(MUTA-) MUTABILIS SA.  
Beaich S;  
WPI; 2004-091381/09.  
DR N-PSDB; ADH88994.  
PT Identifying and selecting a gene required for the proliferation in vivo of a pathogenic microorganism comprises determining the virulence of mutant genes on an experimental model of infection, and their effect on enteric colonization.  
PS Claim 14; SEQ ID NO 13; 60pp; English.

The present invention relates to a method for identifying and selecting a gene required for the proliferation in vivo of a pathogenic microorganism. The method comprises determining the virulence of mutant genes on an experimental model of infection, and their effect on enteric colonization in an axenic mouse model. ADH88967-ADH88981 and ADH88982-ADH88996 are pathogenicity or virulence targets and their coding sequences which were used in the method of the invention. Compounds capable of inhibiting pathogenicity or virulence target expression are useful for developing medicinal products for inhibiting a bacterial infection, in particular an extra-intestinal infection in the case of enterobacteria.

Sequence 477 AA:

Query Match 100.0%; Score 2393; DB 8; Length 477;  
Best Local Similarity 100.0%; Pred. No. 6.9e-207; Mismatches 0; Indels 0; Gaps 0;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKVTLPEPERAGVWVGVDMVMDRMYGPTSRISPEAPVVPVVKVNTIERRPGGANVAMNI 60  
1 MKVTLPEPERAGVWVGVDMVMDRMYGPTSRISPEAPVVPVVKVNTIERRPGGANVAMNI 60  
61 ASIGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRLVSRNOQLRLD 120  
61 ASIGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRLVSRNOQLRLD 120  
121 FEEGFEGVDPQPLHERINQALSSIGALVLSDYAKGALASVQOMIOLARKAGVPLIDPKG 180  
121 FEEGFEGVDPQPLHERINQALSSIGALVLSDYAKGALASVQOMIOLARKAGVPLIDPKG 180  
181 TDFERRYGATLTLPNTSEFAVVGKCKTEBEIYERGMKLIADVELSALVTRSEOGMSLL 240  
181 TDFERRYGATLTLPNTSEFAVVGKCKTEBEIYERGMKLIADVELSALVTRSEOGMSLL 240  
241 QPGKAPLHPMPTQAOBYVDVAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
241 QPGKAPLHPMPTQAOBYVDVAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
301 STVSPTELENAVGRADTGFVMTBEELKLAVAARRGKGVMTNGVFDIILHAGHVSYL 360  
301 STVSPTELENAVGRADTGFVMTBEELKLAVAARRGKGVMTNGVFDIILHAGHVSYL 360  
361 ANARKGDRLLIVAVNSDASTKRLKGSRPVNPLEORMIVGALBAVDVVSFEEDTPORL 420  
361 ANARKGDRLLIVAVNSDASTKRLKGSRPVNPLEORMIVGALBAVDVVSFEEDTPORL 420  
421 IAGILPDLVLVGGDYKPEEIASKEVWANGGEVLVNFEDGCSYTNIIKKIQODKKG 477  
421 IAGILPDLVLVGGDYKPEEIASKEVWANGGEVLVNFEDGCSYTNIIKKIQODKKG 477

RESULT 6  
 ABU28278  
 ID ABU28278 standard; protein; 476 AA.  
 AC ABU28278;  
 DT 19-JUN-2003 (first entry)  
 DE Protein encoded by prokaryotic essential gene #13805.  
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 OS Enterobacter cloacae.  
 XX MO200277183-A2.  
 XX 03-OCT-2002.  
 XX 21-MAR-2002; 2002MO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 XX 06-SEP-2001; 2001US-00948993.  
 XX 25-OCT-2001; 2001US-0342923P.  
 XX 08-FEB-2002; 2002US-00072851.  
 XX 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI, 2003-039926/02.  
 XX N-PSDB; ACA32148.  
 DR New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 25; SEQ ID NO 56202; 1766bp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SO Sequence 476 AA;  
 Query Match 95.9%; Score 2294; DB 6; Length 476;  
 Best Local Similarity 94.7%; Pred. No. 66-198;  
 Matches 451; Conservative 19; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 MKVTLPEFERAGVWVGDVWLDRTYWGPTSRISPEAPVPVVKVNTIEERPGANVMMNI 60  
 DB 1 MKVTLPEFERAGVWVGDVWLDRTYWGPTSRISPEAPVPVVKVNTIEERPGANVMMNI 60  
 QY 61 ASLGANARLVGLTGIDDAARALSKSLADVNVKCDPVSPVPTTITKLRVLSRNOQLRLD 120  
 DB 61 ASLGAGSRVLVGLTGIDDAARALSKSLADVNVKCDPVSPVPTTITKLRVLSRNOQLRLD 120  
 QY 121 FEEGFEVDPPQPIHERINQALSSIGALVSDYAKGALASVQOMIOLARKKGVVLLDPKG 180  
 DB 121 FEEGFEVDPPQPIHERINQALSSIGALVSDYAKGALASVQOMIOLARKKGVVLLDPKG 180  
 QY 121 FEEGFEVDPPQPIHERINQALSSIGALVSDYAKGALASVQOMIOLARKKGVVLLDPKG 180  
 DB 121 FEEGFEVDPPQPIHERINQALSSIGALVSDYAKGALASVQOMIOLARKKGVVLLDPKG 180  
 QY 181 TDPERYRGATLTPNLSPEAVVGVKCTEERIEYERGMKLADYELSLVTRSEOGMSIL 240  
 DB 181 TDPERYRGATLTPNLSPEAVVGVKCTEERIEYERGMKLADYELSLVTRSEOGMSIL 240  
 QY 241 QPKAPLHMPTOAQEYVDVYAGDTVYGVLAATLAAGNSLBEACFPANAAAGVVGKLG 300  
 DB 241 QPKAPLHMPTOAQEYVDVYAGDTVYGVLAATLAAGNSLBEACFPANAAAGVVGKLG 300  
 QY 301 STVSPTELENAVGRADTGFVGWTEBELKLAVAAARKGKRVMTNGVPTILHAGHVSYL 360  
 DB 301 STVSPTELENAVGRADTGFVGWTEBELKLAVAAARKGKRVMTNGVPTILHAGHVSYL 360  
 QY 361 ANARKGDRLLIVVNSDASTKRLKGRPRVNPJEOBMIVYGALEAVDWMVSPEDTPORL 420  
 DB 361 ANARKGDRLLIVVNSDASTKRLKGRPRVNPJEOBMIVYGALEAVDWMVSPEDTPORL 420  
 QY 421 IAGILPDLVKGSDYPRBEIAGSKSEVWANGSEVLYNFEDGCTNNIIRKKIQDDK 476  
 DB 421 IAGILPDLVKGSDYPRBEIAGSKSEVWANGSEVLYNFEDGCTNNIIRKKIQDDK 476  
 RESULT 7  
 ID AAU38166 standard; protein; 477 AA.  
 AC AAU38166;  
 DT 14-FEB-2002 (first entry)  
 DE Salmonella typhi cellular proliferation protein #57.  
 KM Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 OS *Salmonella typhi*.  
 XX MO200170955-A2.  
 XX 27-SEP-2001.  
 XX 21-MAR-2001; 2001MO-US009180.  
 XX 21-MAR-2000; 2000US-0191078P.  
 XX 23-MAY-2000; 2000US-0206848P.  
 XX 26-MAY-2000; 2000US-0207127P.  
 XX 23-OCT-2000; 2000US-0242578P.  
 XX 27-NOV-2000; 2000US-0253625P.  
 XX 22-DEC-2000; 2000US-0257931P.  
 XX 16-FEB-2001; 2001US-0269308P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto R, Xu HH;



Db 1 MKVTLPEFERAGVWVGVMDLRYWVGPTCRISPEAPVPVVKVNTVERPGGANVAMNI 60  
 Qy 61 ASLGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRVLSRNOQLRLD 120  
 Db 61 ASLGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRVLSRNOQLRLD 120  
 Qy 121 FEEGFGVDPQPLHERINQALSGISGALVLSYAKGALTSVQTMIRLAREAGVVLIDPKG 180  
 Db 121 FEEGFGVDPQPLHERINQALSGISGALVLSYAKGALTSVQTMIRLAREAGVVLIDPKG 180  
 Qy 181 TDFERRYGATLTLPNISEFAVAVGKCTEETVERGKMLADVELSALVTRSEOGMSL 240  
 Db 181 TDFERRYGATLTLPNISEFAVAVGKCTEETVERGKMLADVELSALVTRSEOGMSL 240  
 Qy 241 QPGRAPLHPPTQAEVVDVYAGDTVYIGVLAATLAAGNSLEBAECPFANNAAGVVGKLT 300  
 Db 241 QPGRAPLHPPTQAEVVDVYAGDTVYIGVLAATLAAGNSLEBAECPFANNAAGVVGKLT 300  
 Qy 301 STVSPIELBNVAVGRADTGFVMTBEEBLKQAVAAARKGKGVMTNGVPDILAHGVSYL 360  
 Db 301 STVSPIELBNVAVGRADTGFVMTBEEBLKQAVAAARKGKGVMTNGVPDILAHGVSYL 360  
 Qy 361 ANARKGDRILIVAVNSDASTKRLKGSRPVNPLEQRMIVLGALAEADVWVVSFEEDTPQRL 420  
 Db 361 ANARKGDRILIVAVNSDASTKRLKGSRPVNPLEQRMIVLGALAEADVWVVSFEEDTPQRL 420  
 Qy 421 IAGILPDLVYKGGDYKPEBIAGSEKVVANGGEVLVINFEDGCTTNIKKIQODK 476  
 Db 421 IAGILPDLVYKGGDYKPEBIAGSEKVVANGGEVLVINFEDGCTTNIKKIQOTSE 476

RESULT 9  
 ABU32353 ID ABU32353 standard; protein; 477 AA.

XX AC ABU32353;  
 XX DT 19-JUN-2003 (first entry)  
 XX DE Protein encoded by prokaryotic essential gene #17880.  
 XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX OS Klebsiella pneumoniae.  
 XX PN W0200277183-A2.  
 XX PD 03-OCT-2002.  
 XX PF 21-MAR-2002; 2002W0-US009107.  
 XX PR 21-MAR-2001; 2001US-00815242.  
 XX PR 06-SEP-2001; 2001US-00948993.  
 XX PR 25-OCT-2001; 2001US-0342923P.  
 XX PR 08-FEB-2002; 2002US-00072851.  
 XX PR 06-MAR-2002; 2002US-0362699P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX DR WPI, 2003-029926/02.  
 XX DR N-PSDB; ACA36223.  
 XX PT New antisense nucleic acids, useful for identifying proteins or screening  
 XX PT for homologous nucleic acids required for cellular proliferation to  
 XX PT isolate candidate molecules for rational drug discovery programs.  
 XX PS Claim 25, SEQ ID NO 60277; 1766bp; English.  
 XX CC The invention relates to an isolated nucleic acid comprising any one of  
 XX CC the 613 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-regulated gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibody; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIP0 at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SO Sequence 477 AA;

Query Match 93.1%; Score 2227; DB 6; Length 477;

Best Local Similarity 92.2%; Pred. No. 6; 8e-192; Matches 437; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MKVTLPEFERAGVWVGVMDLRYWVGPTCRISPEAPVPVVKVNTVERPGGANVAMNI 60  
 Db 1 MKVTLPEFERAGVWVGVMDLRYWVGPTCRISPEAPVPVVKVNTVERPGGANVAMNI 60  
 Qy 61 ASLGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRVLSRNOQLRLD 120  
 Db 61 ASLGANARLVGLTIGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRVLSRNOQLRLD 120  
 Qy 121 FEEGFGVDPQPLHERINQALSGISGALVLSYAKGALTSVQTMIRLAREAGVVLIDPKG 180  
 Db 121 FEEGFGVDPQPLHERINQALSGISGALVLSYAKGALTSVQTMIRLAREAGVVLIDPKG 180  
 Qy 181 TDFERRYGATLTLPNISEFAVAVGKCTEETVERGKMLADVELSALVTRSEOGMSL 240  
 Db 181 TDFERRYGATLTLPNISEFAVAVGKCTEETVERGKMLADVELSALVTRSEOGMSL 240  
 Qy 241 QPGRAPLHPPTQAEVVDVYAGDTVYIGVLAATLAAGNSLEBAECPFANNAAGVVGKLT 300  
 Db 241 QPGRAPLHPPTQAEVVDVYAGDTVYIGVLAATLAAGNSLEBAECPFANNAAGVVGKLT 300  
 Qy 301 STVSPIELBNVAVGRADTGFVMTBEEBLKQAVAAARKGKGVMTNGVPDILAHGVSYL 360  
 Db 301 STVSPIELBNVAVGRADTGFVMTBEEBLKQAVAAARKGKGVMTNGVPDILAHGVSYL 360  
 Qy 361 ANARKGDRILIVAVNSDASTKRLKGSRPVNPLEQRMIVLGALAEADVWVVSFEEDTPQRL 420  
 Db 361 ANARKGDRILIVAVNSDASTKRLKGSRPVNPLEQRMIVLGALAEADVWVVSFEEDTPQRL 420  
 Qy 421 IAGILPDLVYKGGDYKPEBIAGSEKVVANGGEVLVINFEDGCTTNIKKIQOD 474  
 Db 421 IAGILPDLVYKGGDYKPEBIAGSEKVVANGGEVLVINFEDGCTTNIKKIQOD 474

RESULT 10  
 ABO62691 ID ABO62691 standard; protein; 499 AA.



AC ABO62691;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 DE Klebsiella pneumoniae polypeptide seqid 9208.  
 XX  
 KM Recombinant expression vector; transcription regulatory element;  
 KM Klebsiella pneumoniae protein; antibacterial; Vaccine.  
 XX  
 OS Klebsiella pneumoniae.  
 XX  
 FN US6610836-B1.  
 XX  
 PD 26-AUG-2003.  
 XX  
 PF 27-JAN-2000; 2000US-00489039.  
 XX  
 PR 29-JAN-1999; 99US-0117747P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton GL, Osborne M;  
 XX  
 DR WPI; 2003-895346/82.  
 DR N-PSDB; ACH96242.  
 XX  
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for  
 XX preparing a vaccine composition against Klebsiella pneumoniae.  
 PS Disclosure; SEQ ID NO 9208; 932pp; English.  
 XX  
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella  
 CC pneumoniae polypeptide. Also described are: a recombinant expression  
 CC vector comprising the nucleic acid, operably linked to a transcription  
 CC regulatory element; and a cell comprising the recombinant expression  
 CC vector. The nucleic acid is useful for preparing a vaccine composition  
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a  
 CC Klebsiella pneumoniae polypeptide of the invention  
 XX  
 SQ Sequence 499 AA;  
 Query Match 93.1%; Score 2227; DB 7; Length 499;  
 Best Local Similarity 92.2%; Pred. No. 7.3e-192;  
 Matches 437; Conservative 25; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 MKVTLPFERAGVWVGDVMDRMYGPTSRISPEAPVYVYKVTIERPGGAANVMMNI 60  
 DB 23 MKVTLPFERAGVWVGDVMDRMYGPTSRISPEAPVYVYKVTIERPGGAANVMMNI 82  
 QY 61 ASLGANARLVGLTGIDDAARALSKSLADVNYKDFVSVEPTTITKLRVLSRNQQLRLD 120  
 DB 83 ASLGATSRVLVGLTGIDDAARALSKSLADVNYKDFVSVEPTTITKLRVLSRNQQLRLD 142  
 QY 121 FEEFEGVDPPPLHERINQALSSIGALVLSYAKGALASVQOMQLARKAGVPTLIDPKG 180  
 DB 143 FEEFEGVDPPPLHERINQALSSIGALVLSYAKGALASVQOMQLARKAGVPTLIDPKG 202  
 QY 181 TDFPRYRGATLTTPNLSEFAVWGCKTBEIIVERGMKLIADYELISALVTRSEOGMSLL 240  
 DB 203 TDFPRYRGATLTTPNLSEFAVWGCKTBEIIVERGMKLIADYELISALVTRSEOGMSLL 262  
 QY 241 QPGKAPLHPTQAOEVDVDTGAGDTVIGVLAATLAAGNSLEACFPANAAAAGVVGKLG 300  
 DB 263 QPGKAPLHPTQAOEVDVDTGAGDTVIGVLAATLAAGNSLEACFPANAAAAGVVGKLG 322  
 QY 301 STVPPELEENAVRGRADRGVMTBEELKAAVAARRGEGVNTNGVFDILHAGHSYL 360  
 DB 323 STVPPELEENAVRGRADRGVMTBEELKAAVAARRGEGVNTNGVFDILHAGHSYL 382  
 QY 361 ANAKRLGRLIVAVNSDASTRLKDSRPVNLBOMIVLGALEAVDPAVVSFEEDTPORL 420  
 DB 383 ANAKRLGRLIVAVNSDASTRLKDSRPVNLBOMIVLGALEAVDPAVVSFEEDTPORL 442

QY 421 IAGILPDLVKGGDYKPEIAGSKVWANGSEVLVNFEDGCTNNI IKKI QOD 474  
 DB 443 IAGILPDLVKGGDYKPEIAGSKVWANGSEVLVNFEDGCTNNI IKKI QOD 496  
 RESULT 11  
 ABU45244  
 ID ABU45244 standard; protein; 463 AA.  
 XX  
 AC ABU45244;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by prokaryotic essential gene #30771.  
 XX  
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Salmonella paratyphi.  
 XX  
 FN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362599P.  
 XX  
 PA (BLIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KU, Zykkind JW;  
 PI Wall D, Trawick JD, Carr GF, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA49114.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PS Claim 25; SEQ ID NO 73169; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-regulated gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than S. aureus, S. typhimurium,  
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this



CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 463 AA;  
 SQ Query Match 91.8%; Score 2196; DB 6; Length 463;  
 Best Local Similarity 93.5%; Pred. No. 4,1e-189;  
 Matches 435; Conservative 18; Mismatches 10; Indels 2; Gaps 2;  
 QY 8 FERRAGWVVDVMDRYWYGPTRISPEAPVVPVVKVNTIERPGGANVAMNIALGANA 67  
 DB 1 FERRAGWVVDVMDRYWYGPTRISPEAPVVPVVKVNTIERPGGANVAMNIALGANA 60  
 QY 68 RLVGGLTIDDAARLAKSLADVNVKCPVSVPTHTTKRLVLSRNQOLRLDFEGEG 127  
 DB 61 RLVGGLTIDDAARLAKSLADVNVKCPVSVPTHTTKRLVLSRNQOLRLDFEGEG 119  
 QY 128 VDPQLHERINQALSGALVLSADYAKGALASVQOMIOLARKGVPVLLIDPKGDFERYR 187  
 DB 120 VDPQLHERINQALSGALVLSADYAKGALASVQOMIOLARKGVPVLLIDPKGDFERYR 179  
 QY 188 GATLLTPNLSEFPAVAVKCKTEBEIYERGMKLADYEISALLVTRSEQMSLLQPKAPL 247  
 DB 180 GATLLTPNLSEFPAVAVKCKTEBEIYERGMKLADYEISALLVTRSEQMSLLQPKAPL 239  
 QY 248 HMPPOAEVYDVYAGGTVTGVLATLAAGNSLEBACFPANAAAGVVGKLGSTVSPIE 307  
 DB 240 HMPPOAEVYDVYAGGTVTGVLATLAAGNSLEBACFPANAAAGVVGKLGSTVSPIE 299  
 QY 308 LBNVNRADATGEGVTEBEELKLAVAARKGKGVVMTNGVFDLHAGHVSYLANARKLG 367  
 DB 300 LBNVNRADATGEGVTEBEELKLAVAARKGKGVVMTNGVFDLHAGHVSYLANARKLG 358  
 QY 368 DRLIVANSDASTKRLKGDSPVNPLEQRMIVLGALBAVDVWVVSFEEDTPQRLIAGILPD 427  
 DB 359 DRLIVANSDASTKRLKGDSPVNPLEQRMIVLGALBAVDVWVVSFEEDTPQRLIAGILPD 418  
 QY 428 LLYKGGDYKPEBEIAGSKVWANGGEVLVLFEDDCSTTNIKKIQ 472  
 DB 419 LLYKGGDYKPEBEIAGSKVWANGGEVLVLFEDDCSTTNIKKIQ 463  
 RESULT 12  
 ABUS0241 ID ABUS0241 standard; protein; 476 AA.  
 XX AC ABUS0241;  
 XX DT 19-JUN-2003 (first entry)  
 XX DE Protein encoded by prokaryotic essential gene #35768.  
 XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX OS Versinia pectis.  
 XX PN MO200277183-A2.  
 XX PD 03-OCT-2002.  
 XX PF 21-MAR-2002; 2002MO-US009107.  
 XX PR 21-MAR-2001; 2001US-00815242.  
 XX PR 06-SEP-2001; 2001US-00948993.  
 XX PR 25-OCT-2001; 2001US-0342923P.  
 XX PR 08-FEB-2002; 2002US-00072851.  
 XX PR 06-MAR-2002; 2002US-0362639P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Travick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX WPI; 2003-029926/02.  
 DR N-PSDB; ACAS4111.  
 XX PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PS Claim 25; SEQ ID NO 78165; 1766bp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-regulated gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 476 AA;  
 Query Match 86.8%; Score 2076; DB 6; Length 476;  
 Best Local Similarity 86.1%; Pred. No. 2.9e-178;  
 Matches 409; Conservative 32; Mismatches 34; Indels 0; Gaps 0;  
 QY 1 MKYTLPEFERRAGWVVDVMDRYWYGPTRISPEAPVVPVVKVNTIERPGGANVAMNI 60  
 DB 1 MKYTLPEFERRAGWVVDVMDRYWYGPTRISPEAPVVPVVKVNTIERPGGANVAMNI 60  
 QY 61 ASLIGANARLVGLTGIDDAARLAKSLADVNVKCPVSVPTHTTKRLVLSRNQOLRLD 120  
 DB 61 ASLIGANARLVGLTGIDDAARLAKSLADVNVKCPVSVPTHTTKRLVLSRNQOLRLD 120  
 QY 121 FEEGFGVDPQLHERINQALSGALVLSADYAKGALASVQOMIOLARKGVPVLLIDPKG 180  
 DB 121 FEEGFGVDPQLHERINQALSGALVLSADYAKGALASVQOMIOLARKGVPVLLIDPKG 180  
 QY 181 TDFERRAGATLLTPNLSEFPAVAVKCKTEBEIYERGMKLADYEISALLVTRSEQMSLL 240  
 DB 181 TDFERRAGATLLTPNLSEFPAVAVKCKTEBEIYERGMKLADYEISALLVTRSEQMSLL 240  
 QY 241 QPKAPLHMPPTQAEVYDVYAGGTVTGVLATLAAGNSLEBACFPANAAAGVVGKLG 300  
 DB 241 QPKAPLHMPPTQAEVYDVYAGGTVTGVLATLAAGNSLEBACFPANAAAGVVGKLG 300  
 QY 301 STVSPIELEBANVGRADTGTGVTBEELKLAVAARKGKGVVMTNGVFDLHAGHVSYL 360  
 DB 301 STVSPIELEBANVGRADTGTGVTBEELKLAVAARKGKGVVMTNGVFDLHAGHVSYL 360  
 QY 361 ANARKGDRILIVANSDASTKRLKGDSPVNPLEQRMIVLGALBAVDVWVVSFEEDTPQRL 420

Db 361 ANARKGDRILIVANSDASTKRLKGBKRPVPLEGRVAVVGLALAVDWMVVFEDDTQRL 420  
 Qy 421 IAGILPDLVKGDDYKPEBISKEVWANGGEVLTAVNFBEGCSTTNIKKIQDK 475  
 Db 421 IADILPDLVKGDDYKPEBISKEVWANGGEVLTAVNFBEGCSTTNIKKIQDK 475

## RESULT 13

ID ABU40796 standard; protein; 474 AA.

AC ABU40796;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #26323.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Proteus sp.

PN W0200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX WPI: 2003-029926/02.

DR N-PSDB; ACRA44666.

PT New Antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 68720; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIFO at  
 CC ftp.wifo.int/pub/published\_pct\_sequences

XX Sequence 474 AA;

Query Match 83.9%; Score 2007; DB 6; Length 474;

Best Local Similarity 82.2%; Pred. No. 56-172;

Matches 389; Conservative 43; Mismatches 41; Indels 0; Gaps 0;

Qy 1 MKVTLPEPERAGVWVGVLDVDRWYGPSTRSISPEAPVPVKNVTIEERPGGANVAMNI 60  
 Db 1 MKVTLPEPERAGVWVGVLDVDRWYGPSTRSISPEAPVPVKNVTIEERPGGANVAMNI 60  
 Qy 61 ASIGANARIVGLTIGIDDAARALSKSLADVNVKCDPVSPVTPHTTTLRVLSRNOQLRLD 120  
 Db 61 ASIGANARIVGLTIGIDDAARALSKSLADVNVKCDPVSPVTPHTTTLRVLSRNOQLRLD 120  
 Qy 121 FEERFEGVDPQPHHERINQALSSIGALVSDVAKGALASVQOMIOLARKVGVPLIDPKG 180  
 Db 121 FEERFEGVDPQPHHERINQALSSIGALVSDVAKGALASVQOMIOLARKVGVPLIDPKG 180  
 Qy 181 TDFERYRGATLTPNLSFEFAVVGKCTEIEIVERGKTLADYELSLAVTRSEOGWSLL 240  
 Db 181 TDFERYRGATLTPNLSFEFAVVGKCTEIEIVERGKTLADYELSLAVTRSEOGWSLL 240  
 Qy 241 QPKAPLHMFPTQAEVYDVTGADTYIVGLAATLAAAGNSLSEACFPANAAAGVVGKLTGT 300  
 Db 241 QPKAPLHMFPTQAEVYDVTGADTYIVGLAATLAAAGNSLSEACFPANAAAGVVGKLTGT 300  
 Qy 301 STVSPTELENAVGRADTGVWTEEBELKLAVAARRGKGVMTGVPLILAGHVSYL 360  
 Db 301 STVSPTELENAVGRADTGVWTEEBELKLAVAARRGKGVMTGVPLILAGHVSYL 360  
 Qy 361 ANARKGDRILIVANSDASTKRLKGSRPVPLEGRVAVVGLALAVDWMVVFEDDTQRL 420  
 Db 361 ANARKGDRILIVANSDASTKRLKGSRPVPLEGRVAVVGLALAVDWMVVFEDDTQRL 420  
 Qy 421 IAGILPDLVKGDDYKPEBISKEVWANGGEVLTAVNFBEGCSTTNIKKIQ 475  
 Db 421 IAGILPDLVKGDDYKPEBISKEVWANGGEVLTAVNFBEGCSTTNIKKIQ 475

## RESULT 14

ID ADF06349 standard; protein; 491 AA.

AC ADF06349;

DT 12-FEB-2004 (first entry)

DE Bacterial polypeptide #2462.

XX Proteus mirabilis infection; bacterial infection; antibacterial;

KM immunostimulant.

OS Proteus mirabilis.

PN US6605709-B1.

PD 12-AUG-2003.

PF 05-APR-2000; 2000US-00543681.

PR 09-APR-1999; 99US-0128706P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Breton GL;

DR WPI, 2003-895291/82.  
 DR N-P5DB; ADF02177.  
 PT New Proteus mirabilis polypeptides and polynucleotides, useful as  
 PT reagents for diagnosis of bacterial disease, as components of  
 PT antibacterial vaccines, as targets for antibacterial drugs, or as  
 PT biocontrol agents for plants.  
 XX  
 PS Disclosure; SEQ ID NO 6634; 870bp; English.  
 XX  
 CC The invention relates to new Proteus mirabilis polypeptides and  
 CC polynucleotides. The invention also relates to antibodies against the  
 CC polypeptides, methods for producing the polypeptides, a method of  
 CC generating vaccines for immunising an individual against *P. mirabilis*, a  
 CC method for evaluating a compound for the ability to bind a *P. mirabilis*  
 CC polypeptide and a method for screening test compounds for anti-bacterial  
 CC activity. The polypeptides and polynucleotides are useful as molecular  
 CC targets for diagnosis, preventing and treating pathological conditions  
 CC resulting from bacterial infection, as reagents for diagnosis of  
 CC bacterial diseases, as components of antibacterial vaccines, as targets  
 CC for antibacterial drugs or as bio-control agents for plants. This  
 CC sequence represents a Proteus mirabilis polypeptide of the invention.  
 XX  
 SQ Sequence 491 AA;  
 Query Match 83.9%; Score 2007; DB 7; Length 491;  
 Best Local Similarity 82.2%; Pred. No. 5.2e-172;  
 Matches 389; Conservative 43; Mismatches 41; Indels 0; Gaps 0;  
 QY 1 MKVTLPEERAGVWVGVGDVMDLDRVWGPTSRISPEAPVPVVKNTIERRGAGANVAMNI 60  
 DB 18 MKVTLPPNNKAVLVGDDVMDLDRVWGPTSRISPEAPVPVVKNTIERRGAGANVAMNI 77  
 QY 61 ASIGANARLVGLTGIDDAARLSKSLADVNVKCDPVSVPHTPTTKLRVSRNQLRLD 120  
 DB 78 ASIGANARLVGLTGIDDAARLSKSLADVNVKCDPVSVPHTPTTKLRVSRNQLRLD 137  
 QY 121 PEEFEGVDPPPLHERINQALSSIGALVLSYAKGALASVOQMIQLARKGVPVLIDPKG 180  
 DB 138 PEEFSPNVSPPIYERIQALPSMGALILSYAGALSHVQOMQLAKKGVPLIDPKG 197  
 QY 181 TDFERYGATLLTPNLSFEPAVVGKCTEERIYERGMKLADYELSNLVTREOGMSL 240  
 DB 198 GDFERYGATLLTPNLSFEPAVVGKCTEERIYERGMKLADYELSNLVTREOGMSL 257  
 QY 241 QPGKAPLHMPPTQAEVVDVGTAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLTG 300  
 DB 258 QANBAPLHPTQAEVVDVGTAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLTG 317  
 QY 301 STVSPFIELENNVRGADTGFVMTTEBEIKLAVAAARKGKRVNTNGVPTILAHGVSYL 360  
 DB 318 STVSPFIELENNVRGADTGFVMTTEBEIKLAVAAARKGKRVNTNGVPTILAHGVSYL 377  
 QY 361 ANARKLGDRLIVANSDASTKRLGDSRPVPLBORMIVIGALRAYVMVVSFEEDTPORL 420  
 DB 378 ANARKLGDRLIVANSDASTKRLGDSRPVPLBORMIVIGALRAYVMVVSFEEDTPORL 437  
 QY 421 IAGILPDLVVGDDYKPEEJAGSKREWANGGEVLVNEEDGCTNNIIKKIQ 473  
 DB 438 IASVLPDLVVGDDYKPEEJAGSKREWANGGEVLVNEEDGCTNNIIKKIK 490  
 RESULT 15  
 ABM67735  
 ID ABM67735 standard; protein; 475 AA.  
 AC ABM67735;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Photorhabdus luminescens protein sequence #832.  
 XX  
 KM Antibacterial, fungicide, insecticide, polymorphism, genetic analysis;

KM detection; food; gene expression; plant; animal; microorganism; toxin;  
 KM antibiotic; biopesticide; virulence factor; disease model; plague;  
 KM whooping cough.  
 XX  
 OS Photorhabdus luminescens.  
 XX  
 PN W0200294867-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 07-FEB-2002; 2002MO-IB003040.  
 XX  
 PR 07-FEB-2001; 2001FR-00001659.  
 XX  
 PA (INSP) INST PASTER.  
 PA (CNRS) CNRS CENT NAT RECH SCI.  
 PI Duchaud B, Taourit S, Glaeser P, Frangeul L, Kunat F, Danchin A,  
 PI Buchrieser C;  
 DR WPI, 2003-148459/14.  
 XX  
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides;  
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
 XX  
 PS Claim 2; SEQ ID NO 832; 1205bp; French.  
 XX  
 CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of *P. luminescens*  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than *P. luminescens* and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by *P.*  
 CC luminescens. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically to treat microbial infection by bacteria or fungi that  
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which *P.*  
 CC luminescens is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated *P. luminescens* proteins  
 XX  
 SQ Sequence 475 AA;  
 Query Match 81.5%; Score 1950; DB 6; Length 475;  
 Best Local Similarity 80.6%; Pred. No. 7e-167;  
 Matches 389; Conservative 36; Mismatches 54; Indels 0; Gaps 0;  
 QY 1 MKVTLPEERAGVWVGVGDVMDLDRVWGPTSRISPEAPVPVVKNTIERRGAGANVAMNI 60  
 DB 1 MKVTLPPFCAEVLVGVGDVMDLDRVWGPTSRISPEAPVPVVKNTIERRGAGANVAMNI 60  
 QY 61 ASIGANARLVGLTGIDDAARLSKSLADVNVKCDPVSVPHTPTTKLRVSRNQLRLD 120  
 DB 61 AALGANSRLIGTGIDDAARLSKSLADVNVKCDPVSVPHTPTTKLRVSRNQLRLD 120  
 QY 121 PEEFEGVDPPPLHERINQALSSIGALVLSYAKGALASVOQMIQLARKGVPVLIDPKG 180  
 DB 121 PEEFEGVDVDAQPFERIEQALPHIIGALVLSYAKGALASVOQMIQLARKGVPVLIDPKG 180  
 QY 181 TDFERYGATLLTPNLSFEPAVVGKCTEERIYERGMKLADYELSNLVTREOGMSL 240  
 DB 181 NDFERYGATLLTPNLSFEPAVVGKCTEERIYERGMKLADYELSNLVTREOGMSL 240  
 QY 241 QPGKAPLHMPPTQAEVVDVGTAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLTG 300  
 DB 241 QPGKAPLHMPPTQAEVVDVGTAGDVTIGVLAATLAAGNSLEBACFPANAAAGVVGKLTG 300

Db	241	SVDQPPHLPTQAOBVEFDVTGAGDVTIGVLTATAIAGKPLNEACFLANAAAGVVGKLG	300
Qy	301	STVSPLELENVVRGRADTGFQVMTBEEIKLAVAAARKGEKVMTNGVFDILHAGHVS	360
Db	301	STVSPLELENVVRGRADTGFQVMTBEEIKLAVAAARKGEKVMTNGVFDILHAGHVS	360
Qy	361	ANARKLGDRLLVAVNSDASTYRLKGDSPVNPLEORMIVLGALBAVDWVVSFEEDTPORL	420
Db	361	ENARKLGDRLLVAVNSDASTYRLKGDSPVNPLEORMIVLGALBAVDWVVSFEEDTPORL	420
Qy	421	IAGILPDLLVKGDYKPEBEIAGSKVWANGGEVLVNFEDGCTTNIKKIQDDK	475
Db	421	IADVLPDVLVKGDDYKPEBEIAGSKVWANGGEVLVNFEDGCTTNIKKIQDDK	475

Search completed: March 18, 2006, 06:02:33  
 Job time : 193 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2006, 05:02:48 ; Search time 40 Seconds  
(without alignments)  
1147.385 Million cell updates/sec

Title: US-10-520-820-13

Perfect score: 2393

Sequence: 1 MKVTLPEFERAGVMVGVDMV.....FEDGCTYNIKKIQDDKKG 477

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2393	100.0	477	1 B65093	ADP-heptose synthase
2	2385	99.7	477	2 G91120	ADP-heptose synthase
3	2385	99.7	477	2 F85965	probable heptose kinase Z4
4	2265	94.7	477	2 AG0891	ADP-heptose synthase
5	2076	86.8	476	2 A10080	ADP-heptose synthase
6	1712	71.5	476	1 C64127	ADP-heptose synthase
7	1350.5	56.4	474	2 A83022	LPS biosynthesis P
8	904.5	37.8	483	2 F87700	rfaB protein [impo
9	863	36.1	461	2 H81319	probable ADP-hepto
10	819.5	34.2	463	2 C71887	probable adp-d-gly
11	811	33.9	315	2 G84936	kinase [imported]
12	809.5	33.8	461	1 B64627	ADP-heptose synthase
13	753.5	31.5	323	2 H81154	ADP-heptose synthase
14	745.5	31.2	323	2 H81951	probable DP-heptos
15	594	24.8	463	2 T34841	probable bifunctio
16	564.5	23.6	315	2 H70350	ADP-heptose synthase
17	391.5	16.4	157	2 H70317	glycerol-3-phosphat
18	348.5	14.6	168	2 C81009	aut protein NME207
19	340.5	14.2	168	2 G82031	probable transfe
20	311.5	13.0	164	2 I39548	aut protein - Alca
21	201.5	8.4	319	2 A83882	fructokinase BH185
22	197	8.2	308	2 B83403	fructokinase PA1950
23	194	8.1	299	2 E72311	fructokinase - Therm
24	192.5	8.0	298	2 AH0947	probable sugar kin
25	189.5	7.9	148	2 A75041	probable glycerol-
26	188	7.9	293	1 AB2029	ribokinase [import
27	184.5	7.7	293	1 D69690	ribokinase (EC 2.7
28	184.5	7.7	303	2 B87367	ribokinase [import
29	183.5	7.7	319	2 AC0942	probable carbohydr

30	181.5	7.6	148	2 H71120	hypothetical prote
31	181.5	7.6	378	2 F86307	hypothetical prote
32	179.5	7.5	383	2 D75288	carbohydrate kinas
33	179	7.5	137	2 A69427	glycerol-3-phosphat
34	178	7.4	306	2 B64073	ribokinase (EC 2.7
35	176	7.4	294	2 H84115	ribokinase fbaK [l
36	175	7.3	311	2 D75385	carbohydrate kinas
37	174	7.3	303	2 F84401	hypothetical prote
38	174	7.3	316	2 A96852	fructokinase [impo
39	173.5	7.3	300	2 F91229	probable kinase [l
40	173.5	7.3	300	2 E86076	probable kinase y1
41	170.5	7.1	305	2 T44955	ribokinase (EC 2.7
42	170.5	7.1	343	2 D75260	probable carbohydr
43	170	7.1	300	2 A75599	ribokinase - Deino
44	169	7.1	309	1 K1E3RB	ribokinase (EC 2.7
45	169	7.1	309	2 F91215	ribokinase [import

## ALIGNMENTS

### RESULT 1

B65093

ADP-heptose synthase homolog - Escherichia coli (strain K-12)

N/Alternate names: hypothetical protein b3052

C/Species: Escherichia coli

C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C/Accession: B65093

R/Batner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coj

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; M01D:9742617; EMD:9278503

A/Accession: B65093

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-477 <BLAT>

A/Cross-references: UNIPROT:P76558; UNIPARC:UPI000003EB75; GB:AB000387; GB:U00096; NID:93

A/Experimental source: strain K-12, substrain MG1655

C/Superfamily: hypothetical protein b3052

Query Match	100.0%	Score 2393;	DB 1;	Length 477;
Best local similarity	100.0%	Pred. No. 5.4e-143;		
Matches	477;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;
QY	1	MKVTLPEFERAGVMVGVDMVLDRIWYGPTSRISPEAPVVKVNTTEERPGGAANYAMNI	60	
DB	1	MKVTLPEFERAGVMVGVDMVLDRIWYGPTSRISPEAPVVKVNTTEERPGGAANYAMNI	60	
QY	61	ASIGANARLVGLTGIDDAARALSKSLADVNVKDFVSVPTHTITGLRVLSRQOLRLD	120	
DB	61	ASIGANARLVGLTGIDDAARALSKSLADVNVKDFVSVPTHTITGLRVLSRQOLRLD	120	
QY	121	FEEGFEVDVPPQPIHERINQALSSIGALVLSDYAKGALASVQOMIQARAKGVVLLIDPKG	180	
DB	121	FEEGFEVDVPPQPIHERINQALSSIGALVLSDYAKGALASVQOMIQARAKGVVLLIDPKG	180	
QY	181	TDEERYRGATLLTPNLSPEPAVVGKCTEERIEVERGKLIADYEISALLVTRSEQMSLL	240	
DB	181	TDEERYRGATLLTPNLSPEPAVVGKCTEERIEVERGKLIADYEISALLVTRSEQMSLL	240	
QY	241	QPKAKLHPPTQOQVYVDTGADPTVIGVLAATLAAGNSLEBAQCFPANAAGVVKLG	300	
DB	241	QPKAKLHPPTQOQVYVDTGADPTVIGVLAATLAAGNSLEBAQCFPANAAGVVKLG	300	
QY	301	STSPFIELENAVGRADTGGVWTEBELKLAVAARKEGKVMVTGVPDIIAGHVSYL	360	
DB	301	STSPFIELENAVGRADTGGVWTEBELKLAVAARKEGKVMVTGVPDIIAGHVSYL	360	
QY	361	ANARKIGDRLIIVAVNSDASTKRLKGSRPVNPLEQPMIVIGALBAVDVVSFEEDTPORL	420	
DB	361	ANARKIGDRLIIVAVNSDASTKRLKGSRPVNPLEQPMIVIGALBAVDVVSFEEDTPORL	420	

QY 421 IAGILPDLVKGDDYKKEBIAIGSKKRWANGGEVLVLPNFBDCSTTNIIKKIQDDKKG 477  
 DB 421 IAGILPDLVKGDDYKKEBIAIGSKKRWANGGEVLVLPNFBDCSTTNIIKKIQDDKKG 477

## RESULT 2

G91120  
 ADP-heptose synthase (imported) - Escherichia coli (strain 0157:H7, substrain RMD 05098)  
 C:Species: Escherichia coli  
 C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
 C:Accession: G91120  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
 gasawara, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and gene  
 A:Reference number: A96629; MUID:2156231; PMID:11258796  
 A:Accession: G91120  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-477 <HAV>  
 A:Cross-references: UNIPROT:Q8XBH4; UNIPARC:UPI000000591; GB:BA000007; PIDN:BA037358.1;  
 A:Experimental source: strain 0157:H7, substrain RMD 050952  
 C:Genetic:  
 A:Gene: ECa3935  
 C:Superfamily: hypothetical protein b3052

Query Match 99.7%; Score 2385; DB 2; Length 477;  
 Best Local Similarity 99.6%; Pred. No. 1.7e-142;  
 Matches 475; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVTLPEFERAGVWVGVDMLDRYWYGPTRISPEAPVPVKKVTIEERPGAAVAMNI 60  
 DB 1 MKVTLPEFERAGVWVGVDMLDRYWYGPTRISPEAPVPVKKVTIEERPGAAVAMNI 60  
 QY 61 ASLGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRVLSRNOQLRLD 120  
 DB 61 ASLGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRVLSRNOQLRLD 120  
 QY 121 FEEGFEVDVDPPIHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180  
 DB 121 FEEGFEVDVDPPIHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180  
 QY 181 TDFERYRGATLTLPNLSFEPAVVGKCTEERIVERGKLIADVELSALVTRSEQMSLL 240  
 DB 181 TDFERYRGATLTLPNLSFEPAVVGKCTEERIVERGKLIADVELSALVTRSEQMSLL 240  
 QY 241 QPGKAPLHPMQADEVYDVTAAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
 DB 241 QPGKAPLHPMQADEVYDVTAAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
 QY 301 STVSPIELENAVRGRADTGFVMTSEELKLAVAARRGKGVMTNGVFIDILHAGVSYL 360  
 DB 301 STVSPIELENAVRGRADTGFVMTSEELKLAVAARRGKGVMTNGVFIDILHAGVSYL 360  
 QY 361 ANARKGDRLLIVANSDASTKRLKGSRPVPLEQRMIVIGALBAVDMVVSFEEDTPQRL 420  
 DB 361 ANARKGDRLLIVANSDASTKRLKGSRPVPLEQRMIVIGALBAVDMVVSFEEDTPQRL 420  
 QY 421 IAGILPDLVKGDDYKKEBIAIGSKKRWANGGEVLVLPNFBDCSTTNIIKKIQDDKKG 477  
 DB 421 IAGILPDLVKGDDYKKEBIAIGSKKRWANGGEVLVLPNFBDCSTTNIIKKIQDDKKG 477

## RESULT 3

F85965  
 probable kinase Z4405 (imported) - Escherichia coli (strain 0157:H7, substrain EDL933)  
 C:Species: Escherichia coli  
 C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: F85965  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grocbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potemoules, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F85965  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-477 <STO>  
 A:Cross-references: UNIPROT:Q8XBH4; UNIPARC:UPI000000591; GB:AB005174; NID:G12517628; P  
 A:Experimental source: strain 0157:H7, substrain EDL933  
 C:Genetic:  
 A:Gene: Z4405  
 C:Superfamily: hypothetical protein b3052

Query Match 99.7%; Score 2385; DB 2; Length 477;  
 Best Local Similarity 99.6%; Pred. No. 1.7e-142;  
 Matches 475; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVTLPEFERAGVWVGVDMLDRYWYGPTRISPEAPVPVKKVTIEERPGAAVAMNI 60  
 DB 1 MKVTLPEFERAGVWVGVDMLDRYWYGPTRISPEAPVPVKKVTIEERPGAAVAMNI 60  
 QY 61 ASLGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRVLSRNOQLRLD 120  
 DB 61 ASLGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTTTKLRVLSRNOQLRLD 120  
 QY 121 FEEGFEVDVDPPIHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180  
 DB 121 FEEGFEVDVDPPIHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180  
 QY 181 TDFERYRGATLTLPNLSFEPAVVGKCTEERIVERGKLIADVELSALVTRSEQMSLL 240  
 DB 181 TDFERYRGATLTLPNLSFEPAVVGKCTEERIVERGKLIADVELSALVTRSEQMSLL 240  
 QY 241 QPGKAPLHPMQADEVYDVTAAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
 DB 241 QPGKAPLHPMQADEVYDVTAAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300  
 QY 301 STVSPIELENAVRGRADTGFVMTSEELKLAVAARRGKGVMTNGVFIDILHAGVSYL 360  
 DB 301 STVSPIELENAVRGRADTGFVMTSEELKLAVAARRGKGVMTNGVFIDILHAGVSYL 360  
 QY 361 ANARKGDRLLIVANSDASTKRLKGSRPVPLEQRMIVIGALBAVDMVVSFEEDTPQRL 420  
 DB 361 ANARKGDRLLIVANSDASTKRLKGSRPVPLEQRMIVIGALBAVDMVVSFEEDTPQRL 420  
 QY 421 IAGILPDLVKGDDYKKEBIAIGSKKRWANGGEVLVLPNFBDCSTTNIIKKIQDDKKG 477  
 DB 421 IAGILPDLVKGDDYKKEBIAIGSKKRWANGGEVLVLPNFBDCSTTNIIKKIQDDKKG 477

## RESULT 4

AG0891  
 ADP-heptose synthase (EC 2.7.-.-) (imported) - Salmonella enterica subsp. enterica serov  
 A:Species: Salmonella enterica subsp. enterica serovar Typh  
 A:Note: this species has also been called Salmonella typh  
 C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AG0891  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Conerton, P.; Croft, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Moule, S.; O'Garra, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AG0891  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-477 <PAR>  
 A:Cross-references: UNIPARC:UPI000005A4A3; GB:AL513382; PIDN:CAD07725.1; PID:G16504277; C  
 C:Genetic:  
 A:Gene: rfaB  
 C:Superfamily: hypothetical protein b3052  
 C:Keywords: phosphotransferase

Query Match 94.7%; Score 2265; DB 2; Length 477;



Best Local Similarity 93.5%; Pred. No. 6e-135;  
Matches 445; Conservative 19; Mismatches 12; Indels 0; Gaps 0;

```

Qy 1 MKVTLPEPERAGVAVGVGDMIDRYWGYPTSRISPEAPVYVKNVTIEERPGGAANVANI 60
Db 1 MKVTLPEPERAGVAVGVGDMIDRYWGYPTSRISPEAPVYVKNVTIEERPGGAANVANI 60
Qy 61 ASLGANRLVGLTGIIDDAARALSKSLADVNVKCDPVSPHPTITTKLRVLSRNOQLRLD 120
Db 61 ASLGANRLVGLTGIIDDAARALSKSLADVNVKCDPVSPHPTITTKLRVLSRNOQLRLD 120
Qy 121 FEESFEGVDPPQLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180
Db 121 FEESFEGVDPPQLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180
Qy 181 TDFERYRGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 240
Db 181 TDFERYRGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 240
Qy 241 QPGKAPLHMPTOAQEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 300
Db 241 QPGKAPLHMPTOAQEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 300
Qy 301 STVSPIELENAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 360
Db 301 STVSPIELENAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 360
Qy 361 ANARKDRLIVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 420
Db 361 ANARKDRLIVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 420
Qy 421 IAGILPDLVKGGDYKPEIRIAGSKRWANGSEVYLVNFEDECSSTNIKKIQODK 476
Db 421 IAGILPDLVKGGDYKPEIRIAGSKRWANGSEVYLVNFEDECSSTNIKKIQODK 476

```

## RESULT 5

ADP-heptose synthase (EC 2.7.-.-) [imported] - Yersinia pestis (strain CO92)  
C/Species: Yersinia pestis  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: A10080  
R/Parikhil, J.; Wren, B.W.; Thomson, N.R.; Tibbald, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, W.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001  
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A/Reference number: AB0001; MUID:21470413; PMID:11586360  
A/Accession: A10080  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-476 <TR>  
A/Cross-references: UNIPROT:Q82160; UNIPARC:UPI00000CD706; GB:AL590842; PIDN:CAC89508.1;  
C/Genetics:  
A/Suprafamily: hypothetical protein b3052  
C/Keywords: phosphotransferase

Query Match 86.8%; Score 2076; DB 2; Length 476;  
Best Local Similarity 86.1%; Pred. No. 4.6e-123;  
Matches 409; Conservative 32; Mismatches 34; Indels 0; Gaps 0;

```

Qy 1 MKVTLPEPERAGVAVGVGDMIDRYWGYPTSRISPEAPVYVKNVTIEERPGGAANVANI 60
Db 1 MKVTLPEPERAGVAVGVGDMIDRYWGYPTSRISPEAPVYVKNVTIEERPGGAANVANI 60
Qy 61 ASLGANRLVGLTGIIDDAARALSKSLADVNVKCDPVSPHPTITTKLRVLSRNOQLRLD 120
Db 61 ASLGANRLVGLTGIIDDAARALSKSLADVNVKCDPVSPHPTITTKLRVLSRNOQLRLD 120
Qy 121 FEESFEGVDPPQLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180
Db 121 FEESFEGVDPPQLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180
Qy 181 TDFERYRGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 240
Db 181 TDFERYRGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 240
Qy 241 QPGKAPLHMPTOAQEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 300
Db 241 QPGKAPLHMPTOAQEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 300
Qy 301 STVSPIELENAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 360
Db 301 STVSPIELENAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 360
Qy 361 ANARKDRLIVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 420
Db 361 ANARKDRLIVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 420
Qy 421 IAGILPDLVKGGDYKPEIRIAGSKRWANGSEVYLVNFEDECSSTNIKKIQODK 476
Db 421 IAGILPDLVKGGDYKPEIRIAGSKRWANGSEVYLVNFEDECSSTNIKKIQODK 476

```

```

Qy 181 TDFERYRGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 240
Db 181 TDFERYRGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 240
Qy 241 QPGKAPLHMPTOAQEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 300
Db 241 QPGKAPLHMPTOAQEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 300
Qy 301 STVSPIELENAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 360
Db 301 STVSPIELENAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 360
Qy 361 ANARKDRLIVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 420
Db 361 ANARKDRLIVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 420
Qy 421 IAGILPDLVKGGDYKPEIRIAGSKRWANGSEVYLVNFEDECSSTNIKKIQODK 476
Db 421 IAGILPDLVKGGDYKPEIRIAGSKRWANGSEVYLVNFEDECSSTNIKKIQODK 476

```

## RESULT 6

ADP-heptose synthase homolog - Haemophilus influenzae (strain Rd KW20)  
C/Species: Haemophilus influenzae  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: C64127  
R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Pine, L.D.; Fitchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995  
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A/Reference number: A64000; MUID:95350630; PMID:7542800  
A/Accession: C64127  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-476 <TR>  
A/Cross-references: UNIPROT:O05074; UNIPARC:UPI00001337AB; GB:U32828; GB:L42023; NID:g15;  
C/Suprafamily: hypothetical protein b3052

Query Match 71.5%; Score 1712; DB 1; Length 476;  
Best Local Similarity 71.0%; Pred. No. 3.5e-100;  
Matches 331; Conservative 56; Mismatches 77; Indels 0; Gaps 0;

```

Qy 7 EFERAGVAVGVGDMIDRYWGYPTSRISPEAPVYVKNVTIEERPGGAANVANI 66
Db 7 EFERAGVAVGVGDMIDRYWGYPTSRISPEAPVYVKNVTIEERPGGAANVANI 66
Qy 67 ARVGLTGIIDDAARALSKSLADVNVKCDPVSPHPTITTKLRVLSRNOQLRLD 126
Db 67 ARVGLTGIIDDAARALSKSLADVNVKCDPVSPHPTITTKLRVLSRNOQLRLD 126
Qy 127 GVPQPLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 186
Db 127 GVPQPLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 186
Qy 187 RGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 246
Db 187 RGATLTLPNLSFEPAVNGKCTEERIVERGMKLIADYELSLALVTRSEQMSLL 246
Qy 247 LHMPTQAOEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 306
Db 247 LHMPTQAOEVVDYAGDVTIGVLAATLAAGNSLEAPFANAAAGVVGKLGCT 306
Qy 307 ELKNAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 366
Db 307 ELKNAVRGRADTGFVMTTEBELKLAVAARKGKRVNTNGVFDILHAGHVSYL 366
Qy 367 GDRILVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 426
Db 367 GDRILVAVNSDASTRLKGDSPVNPLEORMITVIGALBAVDVWVSFEEDTPORL 426

```







Db 194 GATLITPBARLKEVGVSMKENDLTEKAQNLRRHLDLTALILTRSEGMTLPSEGR-PI 252  
QY 248 HMPFOAOEVDYTGAGTIVIGVLAATLAAGNSLEBACFPANAAAGVVVGKLGTSVSPIE 307  
Db 253 YQPTFAOEVDVDSGADTVIAGMGLGIAAGCTMPEAMYLANTAGVVAALGTAVCSPAR 312  
QY 308 LENAVRGRA 316  
Db 313 LTKALSGQS 321

## RESULT 15

T34841

probable bifunctional synthase /transferase - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C/Accession: T34841

R/Oliver, K.; Harzif, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, February 1999

A/Reference number: Z21559

A/Accession: T34841

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-463 &lt;OLI&gt;

A/Cross-references: UNIPROT:Q9Z5B5, UNIPARC:UPI00000DAEC2, EMBL:AL035478; PIDN:CAB36595.

A/Experimental source: strain A3(2)

C/Genetic8:

A/Gene: SCODB:SC2G5.08

C/Superfamily: hypothetical protein b3052

Query Match 24.8%; Score 594; DB 2; Length 463;

Best Local Similarity 36.3%; Pred. No. 6.5e-30;

Matches 173; Conservative 71; Mismatches 189; Indels 44; Gaps 13;

QY 9 ERAGVWVGVMDLDRVYGPTRISPEAPVVPVVTIERPGGAANVAMNTASIGANAR 68  
Db 10 DRFPVVVGDALDRDLTGTRADRLAPDAPVPVVGECABRIIRPGALAAVLAARDGREVT 69  
QY 69 LVGLITGIDDAARALSKLADVNVKCDVSVPTHTI-TKLAVLSRNOQLRLDPESGPEG 127  
Db 70 LIAGVGEDPAGLALRELAP---WLKLIPLPLTGTVEKTRVLAQDRPVVRLDRGGG--- 123  
QY 128 VDPQPLHERINQALSSIG---ALVLSYKAGALASVQMIQLAKAGVPLIDPKGTDFE 184  
Db 124 ----RYAEATDEADADGACARAVLVSDYGRGADALRDVL---AARPLVMDPHRGP 175  
QY 185 RYRGATLITNLSEFEAVVGKCKTE-----BEIVERGKMLADYELSLVTRSEGM 237  
Db 176 PVGTTRLVTP--AEKKA-HGFAPSEGRPGGLRAAALNAALVADWRAAVTVTLGSRG- 231  
QY 238 SLLOPGKAPLHMPFOAOEVDYTGAGTIVIGVLAATLAAGNSLEBACFPANAAAGVVVGK 297  
Db 232 ALISYGEHPILVPAAPAHGDSGAGDRFATPAAGLADGALVGEAVEGAVGATAFVAA 291  
QY 298 LGSTVSPILLENVVRGRADTFGVTMBELKLAVAAARKGE--KYVMTNGVFDILHAG 355  
Db 292 GGAANVPASGERALALPTD-----DQALAAIRAEKGVVAAAGCCFDLHAG 342  
QY 356 HVSYLANARKLGDRLIVVNSDASTKRLKDSRPVNPLEQMTVLGALBAVDWVVSPEED 415  
Db 343 HVGILQAARLGDCLVVCNSDASVRKGGKGRPVNPLADRVTLRALACVDAVAVPDED 402  
QY 416 TPQRLIGILPDLVKGDIKPEEI--AGSKVWANGGEVLVNFEDGCTNIKK 470  
Db 403 TPBRLLGELRPDVWVKGDYAGADLPBAGLKEW--GQAVLPLPYLDGRSTALAR 457

Search completed: March 18, 2006, 06:07:13  
Job time : 41 secs

**THIS PAGE LEFT BLANK**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2006, 05:59:28 ; Search time 231 seconds

(without alignments)  
1456.870 Million cell updates/sec

Title: US-10-520-820-13

Perfect score: 2393  
Sequence: 1 MKYTLPEPRAGMVGVDV.....FEDGCTTNLIKXIQDDKGS 477

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2393	100.0	477	1	HLDE_ECOLI
2	2386	99.7	477	1	HLDE_ECOLI
3	2385	99.7	477	1	HLDE_ECOLI
4	2385	99.7	477	1	HLDE_ECOLI
5	2265	94.7	477	1	HLDE_SHIFL
6	2265	94.7	477	1	HLDE_SALTY
7	2265	94.7	477	1	HLDE_SALTY
8	2264	94.6	477	2	OSPC86_SALPA
9	2094	87.5	476	2	Q6D164_ERWCT
10	2081	87.0	476	2	Q65V33_YERPS
11	2076	86.8	476	1	HLDE_YERPS
12	1949	81.4	474	1	HLDE_PHOHL
13	1728	72.2	475	1	HLDE_ACTPL
14	1727	72.2	476	2	Q65T41_MANSN
15	1712	71.5	476	1	HLDE_HASIN
16	1711	71.5	476	2	Q4QXN8_HAB18
17	1701	71.1	476	1	HLDE_PASMU
18	1682	70.3	476	1	HLDE_PHOPR
19	1671	69.8	476	1	HLDE_VIBVU
20	1666	69.6	476	1	HLDE_VIBVY
21	1656	69.2	476	1	HLDE_VIBPA
22	1644	68.7	475	1	HLDE_HABDU
23	1614	67.4	476	2	Q5E2L7_VIBF1
24	1461.5	61.1	476	1	HLDE_SHOON
25	1392	58.2	479	1	HLDE_CAMPB
26	1385.5	57.9	474	2	Q4IUB0_AZOV1
27	1361.5	56.9	474	2	Q4Z215_PSEB1
28	1350.5	56.7	474	1	HLDE_PSEB1
29	1350.5	56.4	474	1	HLDE_PSEB1
30	1350.5	56.4	474	1	HLDE_PSEB1
31	1337.5	55.9	477	2	Q4KJ49_PSEB5

32	1309.5	54.7	473	2	Q607M3_METCA	Q607M3_methylcocc
33	1305	54.5	475	1	HLDE_COXBU	Q3360_coxiella bu
34	955	39.9	490	1	HLDE_GROSL	Q74B6_geobacter s
35	934	39.0	326	2	Q5P2S0_AZOSE	Q5P2S0_azobacter s
36	918	38.4	479	2	Q5F552_GLUOX	Q5F552_gluconobact
37	904.5	37.8	483	1	HLDE_CAUCR	Q9A265_caulobacter
38	891	37.2	470	1	HLDE_WOLSU	Q7M911_wollinella s
39	884.5	37.0	440	1	HLDE_RHOPA	Q6N245_rhodospirillu
40	882.5	36.9	494	1	HLDE_BRAJA	Q6N9R3_bradyrhizob
41	868.5	36.3	457	2	Q4H1T3_CAMDA	Q4H1T3_campylobact
42	865	36.1	482	1	HLDE_HEILAP	Q7V123_helicobacte
43	863	36.1	461	1	HLDE_CAMPJ	Q6G909_campylobact
44	863	36.1	496	1	HLDE_RHIO	Q98154_rhizobium 1
45	855	35.7	461	2	Q5H1W1_CAMPUR	Q5H1W1_campylobact

## ALIGNMENTS

RESULT 1	ID	HLDE_ECOLI	STANDARD;	PRT;	477 AA.
AC	P7658;				
DT	15-JUN-1999	(Rel. 38, Created)			
DT	15-JUN-1999	(Rel. 38, Last sequence update)			
DT	10-MAY-2005	(Rel. 47, Last annotation update)			
DE	Bifunctional protein hldR (includes D-beta-D-heptose 7-phosphate				
DE	kinase (EC 2.7.1.1-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-				
DE	heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)				
GN	Name=hldR; Synonyms=rfab; waag; Ordered locus names=b3052;				
OS	Escherichia coli.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
ON	NCBI_TaxID=562;				
OX	[1]				
RP	NCBIOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].				
RC	STRAIN=K12 / MG1655;				
RX	MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;				
RA	Bacterial F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,				
RA	Riley M., Colado-Vides J., Glaeser J.D., Koe C.K., Mayhew G.F.,				
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,				
RA	Mau B., Shao Y.;				
RT	"The complete genome sequence of Escherichia coli K-12.";				
RT	Science 277:1453-1474(1997).				
RL	[2]				
RP	DOMAINS.				
RC	STRAIN=K12 / W1485;				
RX	MEDLINE=20096694; PubMed=10629197; DOI=10.1128/JB.182.2.488-497.2000;				
RA	Valvano M.A., Marolda C.L., Bittner M., Glaslin-Clay M., Simon T.L.,				
RA	Klena J.D.;				
RT	"The rfa gene from Escherichia coli encodes a bifunctional protein				
RT	involved in the biosynthesis of the lipopolysaccharide core precursor				
RT	ADP-L-glycero-D-manno-heptose.";				
RL	J. Bacteriol. 182:488-497(2000).				
RN	[3]				
RP	FUNCTION. AND ADP-L-BETA-D-HEPTOSE BIOSYNTHESIS PATHWAY.				
RC	STRAIN=K12 / MG1655;				
RX	PubMed=11751812; DOI=10.1128/JB.184.2.363-369.2002;				
RA	Kreidinger B., Marolda C., Graninger M., Zamyatina A., McArthur F.,				
RA	Kosma P., Valvano M.A., Messner P.;				
RT	"Biosynthesis pathway of ADP-L-glycero-beta-D-manno-heptose in				
RT	Escherichia coli.";				
RL	J. Bacteriol. 184:363-369(2002).				
CC	-1- FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-				
CC	heptose 7-phosphate at the C-1 position to form D,D-heptose-1,7-				
CC	bisphosphate.				
CC	-1- FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose				
CC	-1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-				
CC	phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.				
CC	-1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-				
CC	phosphate = diphosphate + ADP-D-glycero-beta-D-manno-heptose.				
CC	-1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-				

```

CC glycerol-beta-D-manno-heptose biosynthesis; second step.
CC -1- PTHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-
CC glycerol-beta-D-manno-heptose biosynthesis; fourth step.
CC -1- MISCELLANEOUS: In both reactions the enzyme functions only with
CC beta anomers.
CC -1- MISCELLANEOUS: The function of the domain II is independent from
CC the activity mediated by domain I.
CC -1- SIMILARITY: In the N-terminal section, belongs to the carbohydrate
CC kinase pfkB family.
CC -1- SIMILARITY: In the C-terminal section, belongs to the
CC cytidylyltransferase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U00096; AAC7608.1; -; Genomic_DNA.
CC PIR; B65093; B65093.
CC HSSP; P27623; 1COZ.
CC Ecobase; E83192; hde.
CC Ecogene; EG13416; hde.
CC HAMAP; MF_01603; -; 1.
CC InterPro; IPR004821; Cyt trans rel.
CC InterPro; IPR004820; Cytidylyltransf.
CC InterPro; IPR002173; pfkB
CC InterPro; IPR011611; pfkB_region.
CC InterPro; IPR011913; RfaE_dom_I.
CC InterPro; IPR011914; RfaE_dom_II.
CC Pfam; PF01467; CTP_transf_2; 1.
CC Pfam; PF00294; pfkB; 1.
CC TIGRfams; TIGR00125; cyt tran rel; 1.
CC TIGRfams; TIGR02198; rfaE_dom_I; 1.
CC TIGRfams; TIGR02199; rfaE_dom_II; 1.
CC PROSITE; PS00583; PFKB_KINASES_1; 1.
CC PROSITE; PS00584; PFKB_KINASES_2; FALSE_NEG.
CC Lipopolysaccharide biosynthesis; Complete proteome; Kinase;
CC KM Nucleotidyltransferase; Multifunctional enzyme;
CC FT REGION 1 318 Ribokinase.
CC REGION 344 477 Cytidylyltransferase.
CC SEQ SEQUENCE 477 AA; 51051 MW; 0F03CB8160B95389 CRC64;

Query March 100.0%; Score 2393; DB 1; Length 477;
Best Local Similarity 100.0%; Pred. No. 2,2e-134;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVTLPFERERAGVWVGVMDLDRYWGFTSRISBEAPVYVKNVTEERPGGANVAMNI 60
DB 1 MKVTLPFERERAGVWVGVMDLDRYWGFTSRISBEAPVYVKNVTEERPGGANVAMNI 60
QY 61 ASLGANARLVGLTGIDDAARLAKSLADVNVKCFVSYPTHTTTTKARVSRNOQLRLD 120
DB 61 ASLGANARLVGLTGIDDAARLAKSLADVNVKCFVSYPTHTTTTKARVSRNOQLRLD 120
QY 121 FEESGEGVDPPLEHERINQALSSIGALVLSDYAGALASVQOMQLARKAGVPVLIIPKG 180
DB 121 FEESGEGVDPPLEHERINQALSSIGALVLSDYAGALASVQOMQLARKAGVPVLIIPKG 180
QY 181 TDFERYRGAATLLTPNLSEFAVVGKCTEERI VERGMKLADYELISALLVTRSEQMSLL 240
DB 181 TDFERYRGAATLLTPNLSEFAVVGKCTEERI VERGMKLADYELISALLVTRSEQMSLL 240
QY 241 QPGKAPLMPLOAQEVYDVYTGAGTVIGVLAATLAAGNSLEBACFPANAAAGVVVGLKGT 300
DB 241 QPGKAPLMPLOAQEVYDVYTGAGTVIGVLAATLAAGNSLEBACFPANAAAGVVVGLKGT 300
QY 301 STVSPPIELNENAVRGADTGFQWMEERELKLVAAARKGKGVYVNTNGVFDLILHAGHSYL 360
DB 301 STVSPPIELNENAVRGADTGFQWMEERELKLVAAARKGKGVYVNTNGVFDLILHAGHSYL 360
QY 361 ANARKLGDRLLIVAVNSDASTKRLKGDSPVNPLEORMIVLGALEAVDWMVVSFEEDTPQRL 420

```

```

DB 361 ANARKLGDRLLIVAVNSDASTKRLKGDSPVNPLEORMIVLGALEAVDWMVVSFEEDTPQRL 420
QY 421 IAGILPDLVKGDDYPEREPIAGSKYEWANGBYLVNPFEDGCTTNIKKIQQDDKXG 477
DB 421 IAGILPDLVKGDDYPEREPIAGSKYEWANGBYLVNPFEDGCTTNIKKIQQDDKXG 477

RESULT 2
HIDE ECOL6
ID HIDE_ECOL6 STANDARD; PRT; 477 AA.
AC Q8FDH5;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Bifunctional protein hde [includes: D-beta-D-heptose 7-phosphate
DE kinase (EC 2.7.1.1-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-
DE heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)];
GN Name=hde; Synonyms=rfaE; OrderedLocustNames=c3800;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:HL / CFT073 / ATCC 700928 / UPEC.
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G., III, Redford P., Reesch P.,
RA Resko D., Buckles E.L., Lion S.R., Boutin A., Hackett J., Stroud D.,
RA Mobley H.L.T., Domeneberg M.S., Blatter F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-
CC heptose 7-phosphate at the C-1 position to form D,D-heptose-1,7-.
CC bisphosphate (By similarity).
CC -1- FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose
CC 1-phosphate, yielding ADP-D,D-heptose (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-
CC phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.
CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-
CC phosphate = ADP + D-glycero-beta-D-manno-heptose.
CC -1- PTHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-
CC glycerol-beta-D-manno-heptose biosynthesis; second step.
CC -1- PTHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-
CC glycerol-beta-D-manno-heptose biosynthesis; fourth step.
CC -1- SIMILARITY: In the N-terminal section, belongs to the carbohydrate
CC kinase pfkB family.
CC -1- SIMILARITY: In the C-terminal section, belongs to the
CC cytidylyltransferase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB016767; AN82245.1; -; Genomic_DNA.
CC HSSP; P27623; 1COZ.
CC HAMAP; MF_01603; -; 1.
CC InterPro; IPR004821; Cyt trans rel.
CC InterPro; IPR004820; Cytidylyltransf.
CC InterPro; IPR002173; pfkB
CC InterPro; IPR011611; pfkB_region.
CC InterPro; IPR011913; RfaE_dom_I.
CC InterPro; IPR011914; RfaE_dom_II.
CC Pfam; PF01467; CTP_transf_2; 1.
CC Pfam; PF00294; pfkB; 1.
CC TIGRfams; TIGR00125; cyt tran rel; 1.
CC TIGRfams; TIGR02198; rfaE_dom_I; 1.
CC TIGRfams; TIGR02199; rfaE_dom_II; 1.
CC PROSITE; PS00583; PFKB_KINASES_1; 1.

```

DR PROSITE; PS00584; PKB KINASES 2; PALSR NEG.  
 KW Carbohydrate metabolism; Complete proteome; Kinase;  
 KW Lipopolysaccharide biosynthesis; Multifunctional enzyme;  
 KW Nucleoside diphosphate kinase.  
 FT REGION 1 318 Ribokinase.  
 FT REGION 344 477 Cytidylyltransferase.  
 SO SEQUENCE 477 AA; 5106 MM; 0P03DBE171495389 CRC64;

Query Match 99.7%; Score 2386; DB 1; Length 477;  
 Best Local Similarity 99.8%; Pred. No. 5.86-134;  
 Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKTLPFERBAGVWVGDVWLDKRYWPTSRISPEAPVPPVKNVTEERPGGANVAMNI 60  
 DB 1 MKTLPFERBAGVWVGDVWLDKRYWPTSRISPEAPVPPVKNVTEERPGGANVAMNI 60

QY 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSPHPITTKLRVLSRNOQLRLD 120  
 DB 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSPHPITTKLRVLSRNOQLRLD 120

QY 121 FEEGFEGVDPPQPIHERINQALSSIGALVLSDYAKGALASVQOMIQARKAGVPLIDPKG 180  
 DB 121 FEEGFEGVDPPQPIHERINQALSSIGALVLSDYAKGALASVQOMIQARKAGVPLIDPKG 180

QY 121 FEEGFEGVDPPQPIHERINQALSSIGALVLSDYAKGALASVQOMIQARKAGVPLIDPKG 180  
 DB 121 FEEGFEGVDPPQPIHERINQALSSIGALVLSDYAKGALASVQOMIQARKAGVPLIDPKG 180

QY 181 TDFERYGATLLTPNLSPEAVVCKTEERIVRGMKLTADYELSLVTRSRQMSLT 240  
 DB 181 TDFERYGATLLTPNLSPEAVVCKTEERIVRGMKLTADYELSLVTRSRQMSLT 240

QY 241 QPGKAPLHPPTQAOEVVDVTGAGDTVIGVLAATLAAGNSLEACFPANAAAGVVGILGT 300  
 DB 241 QPGKAPLHPPTQAOEVVDVTGAGDTVIGVLAATLAAGNSLEACFPANAAAGVVGILGT 300

QY 301 STVSPIELENAVRGRADTGFVMTBEIKALVAARRGKGVVNTNGVFDILAHGHSYL 360  
 DB 301 STVSPIELENAVRGRADTGFVMTBEIKALVAARRGKGVVNTNGVFDILAHGHSYL 360

QY 361 ANAKKLDGRLIVANNSDASTRLKDGSRPNVPLEFOMIVGALFAYVWVVSFEEDTPQL 420  
 DB 361 ANAKKLDGRLIVANNSDASTRLKDGSRPNVPLEFOMIVGALFAYVWVVSFEEDTPQL 420

QY 421 IAGILPDLVKGSDYKPEEIKSGKAVNANGEVVLNPFEDCSYTNIIKKIQDDKKG 477  
 DB 421 IAGILPDLVKGSDYKPEEIKSGKAVNANGEVVLNPFEDCSYTNIIKKIQDDKKG 477

RESULT 3  
 HIDE EC057 STANDARD; PRT; 477 AA.  
 AC 07A07; O8XBW4;  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Bifunctional protein hIdE [includes: D-beta-D-heptose 7-phosphate  
 kinase (EC 2.7.1.-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-  
 heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)];  
 GN Name=hIdE; Synonym=erfAB; OrderedLocName=24405; ECs3935;  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 NX NCBI\_Taxid=83334;  
 RN NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;  
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,  
 Roege D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 Postel G.J., Heckert J., Klink S., Boutin A., Shao Y., Miller L.,  
 Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,  
 Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 Welch R.A., Blattner F.R.;  
 RL "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RN Nature 409:529-533(2001).

RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).  
 RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka N., Tobe T.,  
 Rida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 Kihara S., Shiba T., Hattori M., Shingawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-  
 heptose 7-phosphate (By similarity).  
 CC -1- bisphosphate (By similarity).  
 CC -1- FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose  
 1-phosphate, yielding ADP-D-D-heptose (By similarity).  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-  
 phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-  
 phosphate = diphosphate + ADP-D-glycero-beta-D-manno-heptose.  
 CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-  
 glycerol-beta-D-manno-heptose biosynthesis; second step.  
 CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-  
 glycerol-beta-D-manno-heptose biosynthesis; fourth step.  
 CC -1- SIMILARITY: In the N-terminal section, belongs to the carbohydrate  
 kinase pfkB family.  
 CC -1- SIMILARITY: In the C-terminal section, belongs to the  
 cytidylyltransferase family.  
 CC -----  
 CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use as long as its content is in no way modified and this statement is not  
 removed.  
 CC -----  
 CC EMBL; AB005174; AAC58186.1; -; Genomic DNA.  
 CC EMBL; BA000007; BAB37358.1; -; Genomic DNA.  
 CC PIR; F85965; F85965.  
 CC PIR; G91120; G91120.  
 CC HSP; P27623; 1COZ.  
 CC HAMAP; MF\_01603; -; 1.  
 CC InterPro; IPR004821; Cyt trans rel.  
 CC InterPro; IPR004820; Cytidylyltransf.  
 CC InterPro; IPR002173; pfkB.  
 CC InterPro; IPR011611; pfkB region.  
 CC InterPro; IPR011913; RfaE\_dom\_1.  
 CC InterPro; IPR011914; RfaE\_dom\_1.  
 CC Pfam; PF01467; CTP\_transf\_2; 1.  
 CC Pfam; PF00294; pfkB; 1.  
 CC TIGRFAMs; TIGR00125; Cyt tran rel; 1.  
 CC TIGRFAMs; TIGR02198; rfaE\_dom\_1; 1.  
 CC TIGRFAMs; TIGR02199; rfaE\_dom\_1; 1.  
 CC PROSITE; PS00583; PKB KINASES 1; 1.  
 CC PROSITE; PS00584; PKB KINASES 2; PALSR NEG.  
 KW Carbohydrate metabolism; Complete proteome; Kinase;  
 KW Lipopolysaccharide biosynthesis; Multifunctional enzyme;  
 KW Nucleoside diphosphate kinase.  
 FT REGION 1 318 Ribokinase.  
 FT REGION 344 477 Cytidylyltransferase.  
 SO SEQUENCE 477 AA; 5106 MM; BB877FEB656B67C CRC64;

Query Match 99.7%; Score 2385; DB 1; Length 477;  
 Best Local Similarity 99.6%; Pred. No. 6.66-134;  
 Matches 475; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKTLPFERBAGVWVGDVWLDKRYWPTSRISPEAPVPPVKNVTEERPGGANVAMNI 60  
 DB 1 MKTLPFERBAGVWVGDVWLDKRYWPTSRISPEAPVPPVKNVTEERPGGANVAMNI 60

QY 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSPHPITTKLRVLSRNOQLRLD 120  
 DB 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSPHPITTKLRVLSRNOQLRLD 120

QY 121 FEEGFEGVDPPQPIHERINQALSSIGALVLSDYAKGALASVQOMIQARKAGVPLIDPKG 180

```

Db 121 FESEFEGVDPOPLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180
Qy 181 TDFERRYGATLLTPNLSFEFAVVGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240
Db 181 TDFERRYGATLLTPNLSFEFAVVGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240
Qy 241 QPKRAPLHMTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
Db 241 QPKRAPLHMTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
Qy 301 STVSPIELENAVRGADTGFVMTBEELKLAVAARKGEGVMTNGVPDILAHGHSYL 360
Db 301 STVSPIELENAVRGADTGFVMTBEELKLAVAARKGEGVMTNGVPDILAHGHSYL 360
Qy 361 ANARKLGDRLIVAVNSDASTKRLKGDSPVNPLEORMIVLGALSAVDVMSFEEDTPORL 420
Db 361 ANARKLGDRLIVAVNSDASTKRLKGDSPVNPLEORMIVLGALSAVDVMSFEEDTPORL 420
Qy 421 IAGILPDLVKGDDYKPEEIASKEVWANGEVVLNFEDEGCTTNIKKIQODKXG 477
Db 421 IAGILPDLVKGDDYKPEEIASKEVWANGEVVLNFEDEGCTTNIKKIQODKXG 477

```

```

RESULT 4
HIDE SHIFL STANDARD; PRT; 477 AA.
ID HIDE SHIFL 083047;
AC 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Bifunctional protein hldE (includes: D-beta-D-heptose 7-phosphate
DE kinase (EC 2.7.1.-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-
DE heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)).
GN Name=hldE; Synonyms=rfab; OrderedLocNames=SF3093, S3289;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RA [1]
RA NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RA STRAIN=301 / Serotype 2a;
RA MEDLINE=2227406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang X., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Chang H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RA "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RA through comparison with genomes of Escherichia coli K12 and O157."
RA Nucleic Acids Res. 30:4432-4441 (2002).
RA [2]
RA NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RA STRAIN=2457T / ATCC 700930 / Serotype 2a;
RA MEDLINE=22590274; PubMed=12704152;
RA DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayberg G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RA "Complete genome sequence and comparative genomics of Shigella
RA flexneri serotype 2a strain 2457T."
RA Infect. Immun. 71:2775-2786 (2003).
RA [1]
RA FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-
RA heptose 7-phosphate at the C-1 position to form D,D-heptose-1,7-
RA biphosphate (by similarity).
RA [1]
RA FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose
RA 1-phosphate, yielding ADP-D,D-heptose (by similarity).
RA [1]
RA CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-
RA phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.
RA [1]
RA CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-
RA phosphate = diphosphate + ADP-D-glycero-beta-D-manno-heptose.
RA [1]
RA PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-

```

```

CC glycero-beta-D-manno-heptose biosynthesis; second step.
CC -1 PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-
CC glycero-beta-D-manno-heptose biosynthesis; fourth step.
CC -1 SIMILARITY: In the N-terminal section, belongs to the carbohydrate
CC kinase pfk family.
CC -1 SIMILARITY: In the C-terminal section, belongs to the
CC cytidylyltransferase family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB005674; AAN44569.1; -; Genomic DNA.
CC EMBL; AB016988; AAP18381.1; -; Genomic DNA.
CC HSSP; P27623; 1COZ.
CC DR HMAAP; ME_016031; 1.
CC DR InterPro; IPR004821; Cyt_trans_rel.
CC DR InterPro; IPR004820; Cytidylyltransf.
CC DR InterPro; IPR002173; PfkB.
CC DR InterPro; IPR011611; PfkB_region.
CC DR InterPro; IPR011913; Rfab_dom_1.
CC DR InterPro; IPR011914; Rfab_dom_1.
CC DR InterPro; IPR011914; Rfab_dom_1.
CC DR Pfam; PF01467; CTP_transfer_2; 1.
CC DR Pfam; PF00294; PfkB; 1.
CC DR TIGRFAMs; TIGR00125; Cyt_tran_rel; 1.
CC DR TIGRFAMs; TIGR02198; rfab_dom_1; 1.
CC DR TIGRFAMs; TIGR02199; rfab_dom_1; 1.
CC DR PROSITE; PS00583; PFKB_KINASES_1; 1.
CC DR PROSITE; PS00584; PFKB_KINASES_2; FALSE_NEG.
CC KM Carbohydrate metabolism; Complete proteome; Kinase;
CC KM Lipopolysaccharide biosynthesis; Multifunctional enzyme;
CC KM Nucleotidyltransferase; Transferase.
CC FT REGION 1 318 Rbokinase.
CC FT REGION 344 477 Cytidylyltransferase.
CC FT CONFLCT 133 133 M -> L (in Ref. 2).
CC SQ SEQUENCE 477 AA; 51129 MW; E0F224053402A28 CRC64;
Query Match 99.7%; Score 2385; DB 1; Length 477;
Best Local Similarity 99.6%; Pred. No. 6 6e-134; Indels 0; Gaps 0;
Matches 475; Conservative 1; Mismatches 1;
Qy 1 MKYTLPEFERAGVGVVGVVDVMDRYWGPTSRISPEAPVPVKNVTIEERPGAAVAMNI 60
Db 1 MKYTLPEFERAGVGVVGVVDVMDRYWGPTSRISPEAPVPVKNVTIEERPGAAVAMNI 60
Qy 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSPHPHTTYKLKRVLSRQQLRLD 120
Db 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSPHPHTTYKLKRVLSRQQLRLD 120
Qy 121 FESEFEGVDPOPLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180
Db 121 FESEFEGVDPOPLHERINQALSSIGALVLSDYAKGALASVQOMIQLARKAGVPLIDPKG 180
Qy 181 TDFERRYGATLLTPNLSFEFAVVGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240
Db 181 TDFERRYGATLLTPNLSFEFAVVGKCTEERIVERGKMLADYELSLALVTRSEQMSLL 240
Qy 241 QPKRAPLHMTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
Db 241 QPKRAPLHMTQAOEYVDVDTGAGDTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLG 300
Qy 301 STVSPIELENAVRGADTGFVMTBEELKLAVAARKGEGVMTNGVPDILAHGHSYL 360
Db 301 STVSPIELENAVRGADTGFVMTBEELKLAVAARKGEGVMTNGVPDILAHGHSYL 360
Qy 361 ANARKLGDRLIVAVNSDASTKRLKGDSPVNPLEORMIVLGALSAVDVMSFEEDTPORL 420
Db 361 ANARKLGDRLIVAVNSDASTKRLKGDSPVNPLEORMIVLGALSAVDVMSFEEDTPORL 420
Qy 421 IAGILPDLVKGDDYKPEEIASKEVWANGEVVLNFEDEGCTTNIKKIQODKXG 477
Db 421 IAGILPDLVKGDDYKPEEIASKEVWANGEVVLNFEDEGCTTNIKKIQODKXG 477

```

Db 421 IAGILPDLVKGDPYKPEELIAGSKEVWANGGEVLVLPEDGCSYTTNIIKKIQODKXG 477

RESULT 5

HLDE\_SALTY STANDARD; PRT; 477 AA.

AC OXEN9; 07AM91; (Rel. 45, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DE Bifunctional protein hlde [includes: D-beta-D-heptose 7-phosphate kinase (EC 2.7.1.-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)]

GN Name=hlde; Synonyms=rfab; OrderedLocusNames=STY3339, t3320;

OS Salmonella typhi.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

OC NCBI\_TaxID=601;

RM NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RA STRAIN=CT18;

RA MEDLINE=21531947; PubMed=11677608; DOI=10.1038/35101607;

RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatia M., Baker S., Bauman D., Brooks K., Chillingworth T., Connor P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Keregh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C., Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;

RA "Complete genome sequence of a multiple drug resistant Salmonella RT enterica serovar Typhi CT18."

RT Nature 413:848-852(2001).

RL NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RP STRAIN=Ty2 / ATCC 700931;

RC MEDLINE=2253167; PubMed=12644504;

RX DOI=10.1128/JB.185.7.2330-2337.2003;

RA Deng W., Lion S.-R., Plunkett G. III, Mayhew G.F., Rose D.J., Burland V., Kodyanski V., Schwartz D.C., Blattner F.R.;

RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18."

RL J. Bacteriol. 185:2330-2337(2003).

CC -1- FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-heptose 7-phosphate at the C-1 position to form D,D-heptose-1,7-bisphosphate (By similarity).

CC -1- FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose 1-phosphate, yielding ADP-D,D-heptose (By similarity).

CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.

CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-phosphate = diphosphate + ADP-D-glycero-beta-D-manno-heptose.

CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-glycero-beta-D-manno-heptose biosynthesis; second step.

CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-glycero-beta-D-manno-heptose biosynthesis; fourth step.

CC -1- SIMILARITY: In the N-terminal section; belongs to the carbohydrate kinase pfkB family.

CC -1- SIMILARITY: In the C-terminal section; belongs to the cytidyltransferase family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; AL632728; CAD07725.1; -; Genomic DNA.

CC EMBL; AB016844; AA070663.1; -; Genomic DNA.

CC HSBP; P27623; IC02.

CC HAMAP; MF\_01603; -; 1.

CC InterPro; IPR004821; Cyt trans-rel.

CC InterPro; IPR004820; Cytidylyltransf.

DR InterPro; IPR002173; pfkB.

DR InterPro; IPR011611; pfkB region.

DR InterPro; IPR011913; Rfab\_dom\_1.

DR InterPro; IPR011914; Rfab\_dom\_11.

DR Pfam; PF01467; CTP\_transf\_2; 1.

DR Pfam; PF00294; pfkB; 1.

DR TIGRPFAMs; TIGR00125; cyt tran rel; 1.

DR TIGRPFAMs; TIGR00198; rfab\_dom\_1; 1.

DR TIGRPFAMs; TIGR02199; rfab\_dom\_11; 1.

DR PROSITE; PS00583; PFKB\_KINASES\_1; 1.

DR PROSITE; PS00584; PFKB\_KINASES\_2; FALSE NEG.

KW Carbohydrate metabolism; Complete proteome; Kinase;

KW Lipopolysaccharide biosynthesis; Multifunctional enzyme;

KW Nucleotidyltransferase; Transferase.

FT REGION 1 318 Ribokinase.

FT REGION 344 477 Cytidylyltransferase.

SQ SEQUENCE 477 AA; 51124 MW; E4PF6BDBE80568C CRC64;

Query Match 94.7%; Score 2265; DB 1; Length 477;

Best Local Similarity 93.5%; Pred. No. 9.3e-127;

Matches 445; Conservative 19; Mismatches 12; Indels 0; Gaps 0;

QY 1 MKVTLPEFERAGVYVGVVLDVRYGPTSRISPEAPVPPVKNVTIEERPGAAVAMNT 60

Db 1 MKVNLPAFERAGVWVGVVLDVRYWGPTRISPEAPVPPVKNVTIEERPGAAVAMNI 60

QY 61 ASIGANARLVGLGIDDAARALSKSLADVNVKCDPVSPVPHPTTKLRVSRNOQLRLD 120

Db 61 ASIGANARLVGLGIDDAARALSKTLAEVNVKCDPVSPVPHPTTKLRVSRNOQLRLD 120

QY 121 PEEGPEGVDPPQPHERRINQALSSIGALVLSDYAKGALASVQMIQLARKAGVPLIDPKG 180

Db 121 PEEGPEGVDPPQPHERRINQALSSIGALVLSDYAKGALTSQYMIQLARKAGVPLIDPKG 180

QY 181 TDEERRGATLLTPNISEFAVAVGCKTEBIEVERGKLIADYELSLVTRSEOGQSLL 240

Db 181 TDEERRGATLLTPNISEFAVAVGCKSEBELVERGKLIADYELSLVTRSEOGQNTLL 240

QY 241 QPGKAPLHMTQAOEYVDVDTGADTYIGVLAATLAAGNSLEKCFPNAAGVYVGLTGT 300

Db 241 QPKAPLHMTQAOEYVDVDTGADTYIGVLAATLAAGNTLEBAQYFNAAGVYVGLTGT 300

QY 301 STSPSELENAVGRADPTGCVWTEEBELKLAVAARRGKGVMTNGVFPDILAHGVSYL 360

Db 301 STSPSELENAVGRADPTGCVWTEEBELKQAVASARRGKGVMTNGVFPDILAHGVSYL 360

QY 361 ANARKIGDRLIVAVNSDASTKRLKGSRPVNPLEQRMIVIGALBAVDVWVSFEEDTPQRL 420

Db 361 ANARKIGDRLIVAVNSDASTKRLKGSRPVNPLEQRMIVIGALBSYDVWVSFEEDTPQRL 420

QY 421 IAGILPDLVKGDPYKPEELIAGSKEVWANGGEVLVLPEDGCSYTTNIIKKIQODKX 476

Db 421 IAGILPDLVKGDPYKPEELIAGSKEVWANGGEVLVLPEDGCSYTTNIIKKIQESE 476

RESULT 6

HLDE\_SALTY STANDARD; PRT; 477 AA.

AC Q7CFR9; Q9AJ74; Q9RFY7; Q9RFY8;

DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DE Bifunctional protein hlde [includes: D-beta-D-heptose 7-phosphate kinase (EC 2.7.1.-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)]

GN Name=hlde; Synonyms=rfab; waab; OrderedLocusNames=STY3320;

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

OC NCBI\_TaxID=602;

RM NUCLEOTIDE SEQUENCE, AND ROLE IN LPS BIOSYNTHESIS.

RA PubMed=12441667; DOI=10.1023/A:1021103501626;





DR EMBL/AB017220; AA67053.1; -; Genomic\_DNA.

KM Complete proteome; Transferrase. EAF6B1DBE80566C CRC64;

Seq SEQUENCE 477 AA; 51124 MW; EAF6B1DBE80566C CRC64;  
Query Match 94.7%; Score 2265; DB 2; Length 477;  
Best Local Similarity 93.5%; Pred. No. 9.3e-127;  
Matches 445; Conservative 19; Mismatches 12; Indels 0; Gaps 0;

QY 1 MKVTLPEERAGVNVGDVMDLDRWYGTPTSRISPEAPVPVVKVNTTIERPGGANVAMNI 60  
DB 1 MKVTLPEERAGVNVGDVMDLDRWYGTPTSRISPEAPVPVVKVNTTIERPGGANVAMNI 60  
QY 61 ASLGANARLVGLTGDIDDAARALSKSLADVNVKCPFSVPHTPTTKLRVLSRNOQLRLD 120  
DB 61 ASLGANARLVGLTGDIDDAARALSKSLADVNVKCPFSVPHTPTTKLRVLSRNOQLRLD 120  
QY 121 FEEGFEVDPPQPIHERINQALSGISGALVLSDAKAGALTSVQTMISLARQGVFLIDPKG 180  
DB 121 FEEGFEVDPPQPIHERINQALSGISGALVLSDAKAGALTSVQTMISLARQGVFLIDPKG 180  
QY 181 TDFERRGATLTLPNISEFAVVGKCTEERIVRGKCLADYELSLALVTRSEQMSLL 240  
DB 181 TDFERRGATLTLPNISEFAVVGKCTEERIVRGKCLADYELSLALVTRSEQMSLL 240  
QY 241 QPKAPLHPTQAOEYVDVYAGDPTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
DB 241 QPKAPLHPTQAOEYVDVYAGDPTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
QY 301 STVSPIELBNVNRGRADTGFVMTTEBELKLAVAARRGKGVMTNGVFDILHAGVSYL 360  
DB 301 STVSPIELBNVNRGRADTGFVMTTEBELKLAVAARRGKGVMTNGVFDILHAGVSYL 360  
QY 361 ANARKGDRILIVANSDDASTKRLKGSRPVNPLEQRMIVGALBAVDVWVSFEEDTPQRL 420  
DB 361 ANARKGDRILIVANSDDASTKRLKGSRPVNPLEQRMIVGALBAVDVWVSFEEDTPQRL 420  
QY 421 IAGILPDLVKGDDYKPEEISGKENVANGGEVLVNFEDCGSTTNIKKIQODK 476  
DB 421 IAGILPDLVKGDDYKPEEISGKENVANGGEVLVNFEDCGSTTNIKKIQODE 476

# RESULT 8

Q5PC86\_SALPA PRELIMINARY; PRT; 477 AA.

AC Q5PC86;  
DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE ADP-heptose synthase.  
GN Name=rfar; OrderedLocustNames=SPA3068;  
OS Salmoneella paratyphi-a.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmoneella.  
OX NCBI\_Taxid=54388;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=ATCC 9150;  
RX PubMed=15531882; DOI=10.1038/ng1470;  
RA McCelland M., Sanderson K.E., Clifton S.W., Latreille P.,  
RA Porwollik S., Sabo A., Meyer R., Bierl T., Ozesky P., Mclellan M.,  
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,  
RA Kohlberg S., Strong C., Du F., Carter J., Krenitzki C., Layman D.,  
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Marx P., Florea L.,  
RA Deleahunty K., Fritchard C., Magrini V., Nhan M., Warren W., Florea L.,  
RA Speith J., Wilson R.K.;  
RT Comparison of genome degradation in Paratyphi A and Typhi, human-  
RT restricted serovars of Salmoneella enterica that cause typhoid.";  
RL Nat. Genet. 36:1268-1274(2004).  
DR EMBL/CP000026; AA78903.1; -; Genomic\_DNA.  
DR GO:GO:0016301; P:kinase activity; IEA.  
DR GO:GO:0016779; F:nucleotidyltransferase activity; IEA.  
DR GO:GO:0016740; F:transferase activity; IEA.  
DR GO:GO:0009058; P:biogenesis; IEA.

DR GO:GO:0005975; P:carbohydrate metabolism; IEA.  
DR GO:GO:0009103; P:lipopolysaccharide biosynthesis; IEA.  
DR InterPro:IPR004820; Cytidylyltransferase.  
DR InterPro:IPR004821; Cyt. trans. rel.  
DR InterPro:IPR002173; PfkB.  
DR InterPro:IPR011611; PfkB region.  
DR InterPro:IPR011913; RfaE dom. I.  
DR InterPro:IPR011914; RfaE dom. II.  
DR Pfam:PF01467; CTP. transfer. 2; I.  
DR Pfam:PF0294; PfkB; I.  
DR TIGRPFAMs: TIGR00125; Cyt. trans. rel; 1.  
DR TIGRPFAMs: TIGR02198; rfaE dom. I; 1.  
DR TIGRPFAMs: TIGR02199; rfaE dom. II; 1.  
DR PROSITE: PS00583; PFKB\_KINASES\_1; 1.  
KM Complete proteome.  
SQ SEQUENCE 477 AA; 51138 MW; 3F5C6B1DFC02C69E CRC64;

Query Match 94.6%; Score 2264; DB 2; Length 477;  
Best Local Similarity 93.3%; Pred. No. 1.1e-126;  
Matches 444; Conservative 20; Mismatches 12; Indels 0; Gaps 0;

QY 1 MKVTLPEERAGVNVGDVMDLDRWYGTPTSRISPEAPVPVVKVNTTIERPGGANVAMNI 60  
DB 1 MKVTLPEERAGVNVGDVMDLDRWYGTPTSRISPEAPVPVVKVNTTIERPGGANVAMNI 60  
QY 61 ASLGANARLVGLTGDIDDAARALSKSLADVNVKCPFSVPHTPTTKLRVLSRNOQLRLD 120  
DB 61 ASLGANARLVGLTGDIDDAARALSKSLADVNVKCPFSVPHTPTTKLRVLSRNOQLRLD 120  
QY 121 FEEGFEVDPPQPIHERINQALSGISGALVLSDAKAGALTSVQTMISLARQGVFLIDPKG 180  
DB 121 FEEGFEVDPPQPIHERINQALSGISGALVLSDAKAGALTSVQTMISLARQGVFLIDPKG 180  
QY 181 TDFERRGATLTLPNISEFAVVGKCTEERIVRGKCLADYELSLALVTRSEQMSLL 240  
DB 181 TDFERRGATLTLPNISEFAVVGKCTEERIVRGKCLADYELSLALVTRSEQMSLL 240  
QY 241 QPKAPLHPTQAOEYVDVYAGDPTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
DB 241 QPKAPLHPTQAOEYVDVYAGDPTVIGVLAATLAAGNSLEBACFPANAAAGVVGKLTGT 300  
QY 301 STVSPIELBNVNRGRADTGFVMTTEBELKLAVAARRGKGVMTNGVFDILHAGVSYL 360  
DB 301 STVSPIELBNVNRGRADTGFVMTTEBELKLAVAARRGKGVMTNGVFDILHAGVSYL 360  
QY 361 ANARKGDRILIVANSDDASTKRLKGSRPVNPLEQRMIVGALBAVDVWVSFEEDTPQRL 420  
DB 361 ANARKGDRILIVANSDDASTKRLKGSRPVNPLEQRMIVGALBAVDVWVSFEEDTPQRL 420  
QY 421 IAGILPDLVKGDDYKPEEISGKENVANGGEVLVNFEDCGSTTNIKKIQODK 476  
DB 421 IAGILPDLVKGDDYKPEEISGKENVANGGEVLVNFEDCGSTTNIKKIQODE 476

# RESULT 9

Q6D164\_ERWCT PRELIMINARY; PRT; 476 AA.

AC Q6D164;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE ADP-heptose synthase (EC 2.7.-.-).  
GN Name=rfar; Synonym=vaab; OrderedLocustNames=BCA3584;  
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_Taxid=29471;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=SCRI 1043 / ATCC BAA-672;  
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;  
RA Bell K.S., Seibin M., Pritchard L., Holden M.T.G., Hyman L.J.,  
RA Holve M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,

RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,  
 RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,  
 RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,  
 RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.,  
 RT "Genome sequence of the enterobacterial phytopathogen *Erwinia*  
 RT *carotovora* subsp. *atroseptica* and characterization of virulence  
 RT factors.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).  
 DR EMBL: BX950851; CAG76482.1; -; Genomic DNA.  
 DR GO: GO:0016301; P:kinase activity; IEA.  
 DR GO: GO:0016779; F:nucleotidyltransferase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0009058; P:biosynthesis; IEA.  
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
 DR GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.  
 DR InterPro: IPR004821; Cytidylyltransf.  
 DR InterPro: IPR004821; Cyt\_transf\_rel.  
 DR InterPro: IPR002173; PfkB.  
 DR InterPro: IPR011611; PfkB\_region.  
 DR InterPro: IPR011913; RfaE\_dom\_I.  
 DR InterPro: IPR011914; RfaE\_dom\_II.  
 DR Pfam: PF01467; CTP\_transf\_2; 1.  
 DR Pfam: PF00294; PfkB; 1.  
 DR TIGRFAMs: TIGR00125; Cyt\_tran\_rel; 1.  
 DR TIGRFAMs: TIGR02198; rfaE\_dom\_I; 1.  
 DR TIGRFAMs: TIGR02199; rfaE\_dom\_II; 1.  
 DR PROSITE: PS00583; PFKB\_KINASES\_1; 1.  
 KM Complete proteome; Transferase.  
 SQ SEQUENCE 476 AA; 50763 MW; 19A32257B56B0DB CRC64;

Query Match 87.5%; Score 2094; DB 2; Length 476;  
 Best Local Similarity 87.0%; Pred. No. 1.4e-116;  
 Matches 409; Conservative 30; Mismatches 31; Indels 0; Gaps 0;  
 QY 3 VTLEPEFBAAGVWVGDVWLDRIYWGPTSRISPEAPVPPVVKVNTTIEERFGAANVAMNIAS 62  
 DB 1 MTLTDFRQAGVAVVGDVWLDRIYWGPTSRISPEAPVPPVVKVNTTIEERFGAANVAMNIAA 60  
 QY 63 LGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRVLSRNOQLRLDPE 122  
 DB 61 LGASRLVGLTGVDAAARALAKGAVNVKCDPVSVPHTPTTKLRVLSRNOQLRLDPE 120  
 QY 123 EGFEFVGDVDPQLHERINQALSSIGALVLSDVAKGALASVQOMIQARAKGAVPVLIDPKGT 182  
 DB 121 EGFEFVGDVDPQLHERINQALSSIGALVLSDVAKGALASVQOMIQARAKGAVPVLIDPKGT 180  
 QY 183 FERRRGATLTPNISEFAVVGKCTEERIEYERGMKLIADYELSAALVTRSEQSGMTLLP 242  
 DB 181 FSRRGATLTPNISEFAVVGKCTEERIEYERGMKLIADYELSAALVTRSEQSGMTLLP 240  
 QY 243 GKAPLHMPDQAOEVVDVGTAGDVTYGVAAATLAAGNSLEBACFPANAAAGVVGKLGST 302  
 DB 241 GKAPLHMPDQAOEVVDVGTAGDVTYGVAAATLAAGNSLEBACFPANAAAGVVGKLGST 300  
 QY 303 VSPLEENAVGRADTGVWTEBEELKLAVAARKEGKVVMTNGVPDILHAGVSYLAN 362  
 DB 301 VSPLEENAVGRADTGVWTEBEELKLAVAARKEGKVVMTNGVPDILHAGVSYLAN 360  
 QY 363 ARKAGDRLLIVAVNSDASTKRLKGSRPVNPLEQRNIVGALBAVDWVVSFEEDTPQRL 422  
 DB 361 ARKAGDRLLIVAVNSDASTKRLKGSRPVNPLEQRNIVGALBAVDWVVSFEEDTPQRL 420  
 QY 423 GILPDLVKGDDYKPEEIRAGSEVWANGGEVLVNFEGGCTNNIIKKIQ 472  
 DB 421 GILPDLVKGDDYKPEEIRAGSEVWANGGEVLVNFEGGCTNNIIKKIQ 470

RESULT 10  
 ID 0655V3\_YERPS PRELIMINARY; PRT; 476 AA.  
 AC 0655V3\_2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE ADP-leptose synthase (EC 2.7.-.-).  
 GN Name=rfaf; Synonym=waab; OrderedLocustNames=YPTB3407;  
 OS Yersinia pseudotuberculosis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 ON NCBI\_TaxID=633;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=IP3293 / Serotype I;  
 RX PubMed=15358858; DOI=10.1073/pnas.040401201;  
 RA Chalin P.S.G., Carmel E., Larimer F.W., Lemerding J., Stoutland P.O.,  
 RA Ragala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,  
 RA Rubaker R.R., Fowler J., Himmelsbach J., Marceau M., Medigue C.,  
 RA Simonet M., Chantal-Francoise V., Souza B., Dacheux D., Elliott J.M.,  
 RA Darbise A., Hauser J.J., Garcia E.;  
 RT "Insights into the evolution of *Yersinia* pests through whole-genome  
 RT comparison with *Yersinia pseudotuberculosis*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).  
 DR EMBL: BX36398; CAH2645.1; -; Genomic DNA.  
 DR GO: GO:0016779; F:nucleotidyltransferase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0009058; P:biosynthesis; IEA.  
 DR InterPro: IPR004820; Cytidylyltransf.  
 DR InterPro: IPR004820; Cyt\_transf\_rel.  
 DR InterPro: IPR002173; PfkB.  
 DR InterPro: IPR011611; PfkB\_region.  
 DR InterPro: IPR011913; RfaE\_dom\_I.  
 DR InterPro: IPR011914; RfaE\_dom\_II.  
 DR Pfam: PF01467; CTP\_transf\_2; 1.  
 DR Pfam: PF00294; PfkB; 1.  
 DR TIGRFAMs: TIGR00125; Cyt\_tran\_rel; 1.  
 DR TIGRFAMs: TIGR02198; rfaE\_dom\_I; 1.  
 DR TIGRFAMs: TIGR02199; rfaE\_dom\_II; 1.  
 DR PROSITE: PS00583; PFKB\_KINASES\_1; 1.  
 KM Complete proteome; Transferase.  
 SQ SEQUENCE 476 AA; 51202 MW; B56550CA2D3176BD CRC64;

Query Match 87.0%; Score 2081; DB 2; Length 476;  
 Best Local Similarity 86.3%; Pred. No. 8.5e-116;  
 Matches 410; Conservative 32; Mismatches 33; Indels 0; Gaps 0;  
 QY 1 MKYTLPEFBAAGVWVGDVWLDRIYWGPTSRISPEAPVPPVVKVNTTIEERFGAANVAMNI 60  
 DB 1 MKYTLPEFBAAGVAVVGDVWLDRIYWGPTSRISPEAPVPPVVKVNTTIEERFGAANVAMNI 60  
 QY 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRVLSRNOQLRLD 120  
 DB 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRVLSRNOQLRLD 120  
 QY 121 FEEGFEGVDVDPQLHERINQALSSIGALVLSDVAKGALASVQOMIQARAKGAVPVLIDPKG 180  
 DB 121 FEEGFEGVDVDPQLHERINQALSSIGALVLSDVAKGALASVQOMIQARAKGAVPVLIDPKG 180  
 QY 181 TDERRRGATLTPNISEFAVVGKCTEERIEYERGMKLIADYELSAALVTRSEQSGMTLL 240  
 DB 181 SDEERYGATLTPNISEFAVVGKCTEERIEYERGMKLIADYELSAALVTRSEQSGMTLL 240  
 QY 241 QPQKAPLHMPDQAOEVVDVGTAGDVTYGVAAATLAAGNSLEBACFPANAAAGVVGKLGST 300  
 DB 241 QPQKAPLHMPDQAOEVVDVGTAGDVTYGVAAATLAAGNSLEBACFPANAAAGVVGKLGST 300  
 QY 301 STVSPLEENAVGRADTGVWTEBEELKLAVAARKEGKVVMTNGVPDILHAGVSYLAN 360  
 DB 301 STVSPLEENAVGRADTGVWTEBEELKLAVAARKEGKVVMTNGVPDILHAGVSYLAN 360  
 QY 361 ANARKIGDRLLIVAVNSDASTKRLKGSRPVNPLEQRNIVGALBAVDWVVSFEEDTPQRL 420  
 DB 361 ANARKIGDRLLIVAVNSDASTKRLKGSRPVNPLEQRNIVGALBAVDWVVSFEEDTPQRL 420  
 QY 421 IAGILPDLVKGDDYKPEEIRAGSEVWANGGEVLVNFEGGCTNNIIKKIQ 475  
 DB 421 IADILPDLVKGDDYKPEEIRAGSEVWANGGEVLVNFEGGCTNNIIKKIQ 475



DT 25-OCT-2004 (Rel. 45, last sequence update)  
 DT 10-MAY-2005 (Rel. 47, last annotation update)  
 DE Bifunctional protein hldR [includes: D-beta-D-heptose 7-phosphate  
 DE kinase (EC 2.7.1.-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-  
 DE heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)].  
 GN Name=hldR; Synonyms=rfas; OrderedLocNames=plu3968;  
 OS Photobacterium luminescens (subsp. laumondii).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photobacterius.  
 NCBI\_TaxID=141679;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=TT01;  
 RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;  
 RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,  
 RA Taourit S., Bocs S., Bouraux-Eude C., Chandler M., Charles J.-F.,  
 RA Dassa E., Deroose R., Derzelle S., Freysinet G., Gaudriault S.,  
 RA Meligne C., Lantols A., Powell K., Sigulier P., Vincent R., Wingate V.,  
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;  
 RT "The genome sequence of the entomopathogenic bacterium Photobacterius  
 RT luminescens";  
 RT Nat. Biotechnol. 21:1307-1313(2003).  
 CC -1- FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-  
 CC heptose 7-phosphate at the C-1 position to form D,D-heptose-1,7-  
 CC biphosphate (By similarity).  
 CC -1- FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose  
 CC 1-phosphate, yielding ADP-D,D-heptose (By similarity).  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-  
 CC phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-  
 CC phosphate = diphosphate + ADP-D-glycero-beta-D-manno-heptose.  
 CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-  
 CC glycero-beta-D-manno-heptose biosynthesis; second step.  
 CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-  
 CC glycero-beta-D-manno-heptose biosynthesis; fourth step.  
 CC -1- SIMILARITY: In the N-terminal section; belongs to the carbohydrate  
 CC kinase pfkB family.  
 CC -1- SIMILARITY: In the C-terminal section; belongs to the  
 CC cytidylyltransferase family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL: BX571872; CAE16340.1; -; Genomic\_DNA.  
 CC Photobacterius plu3968; -;  
 DR HAMAP: MF\_01603; -; 1.  
 DR InterPro: IPR004821; Cyt trans rel.  
 DR InterPro: IPR004820; Cytidylyltransf.  
 DR InterPro: IPR002173; PEKB.  
 DR InterPro: IPR011611; PEKB region.  
 DR InterPro: IPR011913; RfaE\_dom\_1.  
 DR InterPro: IPR011914; RfaE\_dom\_11.  
 DR Pfam: PF01467; CYP transf\_2; 1.  
 DR Pfam: PF00294; PEKB; 1.  
 DR TIGRfams: TIGR00125; Cyt tran rel; 1.  
 DR TIGRfams: TIGR02198; rfaE\_dom\_1; 1.  
 DR TIGRfams: TIGR02199; rfaE\_dom\_11; 1.  
 DR TIGRfams: TIGR02199; rfaE\_dom\_11; 1.  
 DR PROSITE: PS00583; PEKB KINASES\_1; 1.  
 DR PROSITE: PS00584; PEKB KINASES\_2; FALSE NEG.  
 KM Carbohydrate metabolism; Complete proteome; Kinase;  
 KM Lipopolysaccharide biosynthesis; Multifunctional enzyme;  
 KM Nucleotidyltransferase; Transferase.  
 FT REGION 1 318 Ribokinase.  
 FT REGION 344 474 Cytidylyltransferase.  
 SQ SEQUENCE 474 AA; 5130 MW; DCAC4E3DD108862D CRC64;  
 Query Match 81.4%; Score 1949; DB 1; Length 474;  
 Best Local Similarity 81.1%; Pred. No. 6.2e-108;  
 Matches 383; Conservative 37; Mismatches 52; Indels 0; Gaps 0;

Qy 1 MKVTLDEFERAGVGVGVVMDLRYWYGPSTRISPEAPVPVVKVNTIIEERGAANVAMNI 60  
 Db 1 MKVTLDPFHCABVLYVGVVDLDRYWGPTSRISPEAPVPVVKVNTIIEERGAANVAMNI 60  
 Qy 61 ASIGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPFHPTTTLKRVTSRQQLRLD 120  
 Db 61 AALGANSHLLGLGIDDAARALSEKRSVYVRCDFPSVPFHPVTYKLRVTSRQQLRLD 120  
 Qy 121 PEEGPEGVDPPQPHERRINQALSSIGALVLSDYAKGALASVQMIQLARKAGVPLIDPKG 180  
 Db 121 PEEGFDVNDQPFEPHIEQLPHITGALVLSDYAKGALASVQMIKLANAKVVLIDPKG 180  
 Qy 181 TDFERRYGATLLPNTSEFAVVGKCTEEBIEVRGKLLADYBSALLVTSRBOGMSLL 240  
 Db 181 NDFERRYGATLLPNTSEFAVVGKCDDELVEKGRVLYKGDQLQALLITRSEOGMSLL 240  
 Qy 241 QPKGAPLHMTQOEYVDVTGADVTYIGVLAATLAAAGNSLEBECFPANAAAGVVGGLCT 300  
 Db 241 SVDPPLHLPTQOEYVDVTGADVTYIGVLAATLAAAGNSLEBECFPANAAAGVVGGLGT 300  
 Qy 301 STVSPTELENAVGRADTGFVMTTEBELKLAVAARKRGKVMVTNGVPDILAHGVSYL 360  
 Db 301 STVSPTELENAITGRATGRTGVTTEFOLKQAVDARQGRVMTNCGPIILAHGVSYL 360  
 Qy 361 ANARKGDRLLIVAVNSDASTKRLKGSRPVNPLEQMTVYLALEAVDVVSFEEDTPQRL 420  
 Db 361 ENARRIGDRLLIVAVNSDASTKRLKGSRPVNPLEQMTVYLALEAVDVVSFEEDTPQRL 420  
 Qy 421 IAILPDLVLYKGGDYPEREIRAGSEKYEYVANGGEVLVNFEGCGSTYNIKKIQ 472  
 Db 421 IADVLPLVLYKGGDYPEREIRAGSEYVANGGEVLVNFEGISTYNIKAIR 472  
 RESULT 13  
 HIDE ACTPL  
 ID HIDE ACTPL STANDARD; PRT; 475 AA.  
 AC QGGLDY;  
 DT 25-OCT-2004 (Rel. 45, last sequence update)  
 DT 10-MAY-2005 (Rel. 47, last annotation update)  
 DE Bifunctional protein hldR [includes: D-beta-D-heptose 7-phosphate  
 DE kinase (EC 2.7.1.-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-  
 DE heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)].  
 GN Name=hldR; Synonyms=rfas;  
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Actinobacillus.  
 NCBI\_TaxID=715;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE (GENOMIC DNA), AND ROLE IN LPS BIOSYNTHESIS.  
 RC STRAIN=S 4074 / Serotype 1;  
 RX MEDLINE=22683163; PubMed=12798993; DOI=10.1016/S0378-1097(03)00247-7;  
 RA Provost M., Harel J., Labrie J., Sirolis M., Jacques M.;  
 RT "Identification, cloning and characterization of rfbE of  
 RT Actinobacillus pleuropneumoniae serotype 1, a gene involved in  
 RT lipopolysaccharide inner-core biosynthesis";  
 RT FEMS Microbiol. Lett. 223:7-14(2003).  
 RL FEMS Microbiol. Lett. 223:7-14(2003).  
 CC -1- FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-  
 CC heptose 7-phosphate at the C-1 position to form D,D-heptose-1,7-  
 CC biphosphate (Probable).  
 CC -1- FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose  
 CC 1-phosphate, yielding ADP-D,D-heptose (Probable).  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-  
 CC phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-  
 CC phosphate = diphosphate + ADP-D-glycero-beta-D-manno-heptose.  
 CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-  
 CC glycero-beta-D-manno-heptose biosynthesis; second step.  
 CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis; ADP-L-  
 CC glycero-beta-D-manno-heptose biosynthesis; fourth step.  
 CC -1- SIMILARITY: In the N-terminal section; belongs to the carbohydrate  
 CC kinase pfkB family.  
 CC -1- SIMILARITY: In the C-terminal section; belongs to the



DT 15-JUN-1999 (Rel. 38, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Bifunctional protein htdB [includes: D-beta-D-heptose 7-phosphate  
 DE kinase (EC 2.7.1.-) (D-beta-D-heptose 7-phosphotransferase); D-beta-D-  
 DE heptose 1-phosphate adenosyltransferase (EC 2.7.7.-)]  
 GN Name=1hde; Synonyms=rtab, waab; OrderedLocustNames=H11526;  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OC NCBI\_TaxId=727;  
 RX [1]  
 RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RP STRAIN=rd / KW20 / ATCC 51907;  
 RC MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirtness E.F.,  
 RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geophagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE OF 1-427, AND ROLE IN LOS BIOSYNTHESIS.  
 RC STRAIN=NTM1 2019;  
 RX MEDLINE=95172727; PubMed=7868252;  
 RT Lee N.-G., Sunshine M.G., Apicella M.A.;  
 RT "Molecular cloning and characterization of the nontypeable Haemophilus  
 RT influenzae 2019 rfaE gene required for lipopolysaccharide  
 RT biosynthesis.";  
 RL Infect. Immun. 63:818-824(1995).  
 CC -1- FUNCTION: Catalyzes the phosphorylation of D-glycero-D-manno-  
 CC heptose 7-phosphate at the C-1 position to form D,D-heptose-1,7-  
 CC bisphosphate (Probable).  
 CC -1- FUNCTION: Catalyzes the ADP transfer to D-glycero-D-manno-heptose  
 CC 1-phosphate, yielding ADP-D,D-heptose (Probable).  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 7-  
 CC phosphate = ADP + D-glycero-beta-D-manno-heptose 1,7-bisphosphate.  
 CC -1- CATALYTIC ACTIVITY: ATP + D-glycero-beta-D-manno-heptose 1-  
 CC phosphate = diphosphate + ADP-D-glycero-beta-D-manno-heptose.  
 CC -1- PATHWAY: Lipopoligosaccharide (LOS) core biosynthesis; ADP-L-  
 CC glycero-beta-D-manno-heptose biosynthesis; second step.  
 CC -1- PATHWAY: Lipopoligosaccharide (LOS) core biosynthesis; ADP-L-  
 CC glycero-beta-D-manno-heptose biosynthesis; fourth step.  
 CC -1- SIMILARITY: In the N-terminal section; belongs to the carbohydrate  
 CC kinase pfkB family.  
 CC -1- SIMILARITY: In the C-terminal section; belongs to the  
 CC cytidylyltransferase family.  
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 331.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL: U32828; AAC23172.1; -; Genomic DNA.  
 DR EMBL: U17642; AAC43516.1; ALT\_FRAME; Genomic DNA.  
 DR PIR: C64127; C64127.  
 DR HSPB; P27623; 1COZ.  
 DR TIGR: H11526; -.  
 DR HAMAP: MF\_01603; -; 1.  
 DR InterPro: IPR004821; Cyt trans rel.  
 DR InterPro: IPR004820; Cytidylyltransf.  
 DR InterPro: IPR002173; pfkB.  
 DR InterPro: IPR011611; pfkB\_region.  
 DR InterPro: IPR011913; RfaE\_dom\_1.  
 DR InterPro: IPR011914; RfaE\_dom\_1i.

DR Pfam; PF01467; CTP transf\_2; 1.  
 DR Pfam; PF00294; pfkB; 1.  
 DR TIGRPFam; TIGR00125; Cyt tran rel; 1.  
 DR TIGRPFam; TIGR02198; rfaE\_dom\_1; 1.  
 DR TIGRPFam; TIGR02199; rfaE\_dom\_1i; 1.  
 DR PROSITE; PS00583; pfkB\_KINASES\_1; FALSE\_NEG.  
 DR PROSITE; PS00584; pfkB\_KINASES\_2; FALSE\_NEG.  
 DR Carbohydrate metabolism; Complete proteome; Kinase;  
 KM Multifunctional enzyme; Nucleotidylyltransferase; Transferase.  
 FT REGION 1 318  
 FT REGION 344 476  
 FT CONFLICT 36 36 Cytidylyltransferase.  
 FT CONFLICT 84 84 A -> R (in Ref. 2).  
 FT CONFLICT 170 170 L -> H (in Ref. 2).  
 FT CONFLICT 170 170 A -> G (in Ref. 2).  
 FT CONFLICT 214 214 E -> K (in Ref. 2).  
 FT CONFLICT 273 273 A -> T (in Ref. 2).  
 FT CONFLICT 278 278 R -> C (in Ref. 2).  
 SQ SEQUENCE 476 AA; 51946 MW; 4F241C08D2C6951E CRC64;  
 Query Match 71.5%; Score 1712; DB 1; Length 476;  
 Best local Similarity 71.0%; Pred. No. 8,2e-94;  
 Matches 331; Conservative 58; Mismatches 77; Indels 0; Gaps 0;  
 QY 7 EFERAGVWVGVDMVDRVYGTPTSRISPEAPVPVKVNTIEEPGGANVANNIASIGAN 66  
 DB 7 EFKQAKVLVGVDMVDRVYGTPTSRISPEAPVPVKVNTIEEPGGANVANNIASINVP 66  
 QY 67 ARVLGLTGIDDAARALSKSLADVNVKDFVSPVTHPTITRLVRSNQQLRIIDFEEGFE 126  
 DB 67 VQLMGILGDETSASLLEKOKIDCNFVALTHPTITRLVRSNQQLRIIDFEEDEFN 126  
 QY 127 GVPOPLHERINOALSSIGALVSDYAKGALASVQOMIQLARKAGVVLIDPGKTFERY 186  
 DB 127 NVCKKLLATLSAENVYKGLISDYGKTLKVOGCMIGIARKANVPVILDPKGTDERY 186  
 QY 187 RGATLTLPNLSFEFAVVGKCTEEBIVERGMKLIADYEISALIVTSSEQMSILQPEKAP 246  
 DB 187 RGATLTLPNMSFEFAVVGKCTEEBIEIKGLKISIEILTALVTSSEKMTLIRPVQEP 246  
 QY 247 LHMPTQAOEVYDVTAGDPTVIGVLAATTLAAGNLEBACFPANNAAGVVGKLTSTVSP 306  
 DB 247 YHLPTVAKEVFYDVTAGDPTVISVLAATLADGRFBESCYLANAAGVVGKLTSTVSTV 306  
 QY 307 ELENAVGRADTGFVMTTEBELKLAVALARKKRGKVTMTNGVDPDILHAGVSYLANARKL 366  
 DB 307 ELENAIHARETFGIMSEBELKDAVAQAKAREKIVMTNGVCPDILHPGISYLENARKL 366  
 QY 367 GDRILVAVNSDASTKRLKGDSPVNPLEQRMIVLGALEAVDVVVSFEEDTPORLIAGILP 426  
 DB 367 GDRILVAVNSDDSVKRLKGSREPINLLENMAVLAGIASVDMLVPFTEDTPORLIGELIP 426  
 QY 427 DLLVKGDDYKEERIASKKEVWANGGVLYLNFEDGSTTNIKKIQ 472  
 DB 427 DLLVKGDDYKEERIASKKEVWANGGVLYLNFENGSTTNIKKIK 472

Search completed: March 18, 2006, 06:06:28  
 Job time : 234 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2006, 06:06:43, Search time 48 Seconds  
(without alignments)  
821.590 Million cell updates/sec

Title: US-10-520-820-13

Profile score: 2393

Sequence: 1 MKTTLPEFBAAGVWVGDVW.....PBDGCTTNIKKIQDQDKG 477

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:\*

1: /cgn2\_6/ptcdat/1/1aa/5\_COMB.pep:\*

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2393	100.0	477	US-09-492-709A-325	Sequence 325, App
2	2227	93.1	499	US-09-489-039A-9208	Sequence 9208, Ap
3	2007	83.9	491	US-09-543-681A-6634	Sequence 6634, Ap
4	1351.5	56.5	526	US-09-252-991A-32984	Sequence 32984, A
5	610.5	25.5	498	US-09-477-962-114	Sequence 114, App
6	197	8.2	342	US-09-489-039A-9668	Sequence 9668, Ap
7	190.5	8.2	503	US-09-252-991A-31882	Sequence 31882, A
8	190.5	8.0	529	US-09-134-000C-4524	Sequence 4524, Ap
9	183	7.6	311	US-09-543-681A-7498	Sequence 7498, Ap
10	171	7.1	356	US-09-489-039A-12757	Sequence 12757, A
11	169	7.1	423	US-09-668-262A-14	Sequence 14, App1
12	169	7.1	423	US-10-427-442-14	Sequence 14, App1
13	167.5	7.0	665	US-09-489-039A-8800	Sequence 8800, Ap
14	165.5	6.9	315	US-09-134-001C-4521	Sequence 4521, Ap
15	165	6.9	330	US-09-489-039A-14043	Sequence 14043, A
16	163.5	6.8	421	US-09-668-262A-16	Sequence 16, App1
17	163.5	6.8	421	US-10-427-442-16	Sequence 16, App1
18	155	6.5	320	US-09-543-681A-5963	Sequence 5963, Ap
19	154.5	6.3	302	US-09-710-279-3062	Sequence 3062, Ap
20	150	6.3	137	US-09-134-001C-5560	Sequence 5560, Ap
21	146.5	6.1	269	US-09-270-767-6416	Sequence 6416, A
22	146	6.1	208	US-09-270-767-61995	Sequence 61995, A
23	143	6.0	293	US-09-668-262A-12	Sequence 12, App1
24	143	6.0	293	US-10-427-442-12	Sequence 12, App1
25	143	6.0	311	US-09-543-681A-5777	Sequence 5777, Ap
26	141	5.9	313	US-09-489-039A-12895	Sequence 12895, A
27	141	5.9	355	US-09-248-796A-16878	Sequence 16878, A

28	139.5	5.8	331	2	US-09-902-540-10221	Sequence 10221, A
29	137.5	5.7	312	2	US-09-134-001C-5583	Sequence 5583, Ap
30	137.5	5.7	347	2	US-08-826-611-2	Sequence 2, App1
31	134.5	5.6	3724	1	US-08-804-227C-10	Sequence 10, App1
32	134.5	5.6	3724	1	US-08-804-198-4	Sequence 4, App1
33	133.5	5.6	308	2	US-08-913-816C-17	Sequence 17, App1
34	131	5.5	143	2	US-09-134-000C-4544	Sequence 4544, Ap
35	130	5.4	859	2	US-09-199-637A-281	Sequence 281, App
36	127	5.3	859	2	US-09-252-991A-21413	Sequence 21413, A
37	125	5.2	314	2	US-09-489-039A-10377	Sequence 10377, A
38	125	5.2	331	2	US-09-543-681A-7111	Sequence 7111, Ap
39	123.5	5.2	328	2	US-09-489-039A-7290	Sequence 7290, Ap
40	121	5.1	301	2	US-09-252-991A-22565	Sequence 22565, A
41	121	5.1	406	2	US-09-543-681A-8124	Sequence 8124, Ap
42	120.5	5.0	269	2	US-09-134-000C-5214	Sequence 5214, Ap
43	119.5	5.0	411	2	US-09-949-016-8628	Sequence 8628, Ap
44	118	4.9	595	2	US-09-902-540-16334	Sequence 16334, A
45	118	4.9	719	2	US-09-902-540-11526	Sequence 11526, A

## ALIGNMENTS

RESULT 1

US-09-492-709A-325

Sequence 325, Application US/09492709A

Patent No. 6720139

GENERAL INFORMATION:

APPLICANT: Zykend, Judith

APPLICANT: Onlsen, Karl L.

APPLICANT: Treawick, John

APPLICANT: Forsyth, R. Allyn

APPLICANT: Froelich, Jamie M.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN

FILE REFERENCE: ELITRA.001A

CURRENT APPLICATION NUMBER: US/09/492,709A

CURRENT FILING DATE: 2000-01-27

NUMBER OF SEQ ID NOS: 485

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 325

LENGTH: 477

TYPE: PRT

ORGANISM: E. Coli

US-09-492-709A-325

Query Match

Best Local Similarity 100.0%; Score 2393; DB 2; Length 477;

Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKTTLPEFBAAGVWVGDVWLDRTYWGPTSRISPEAPVPVKNVTIEERPGANVAMNT	60
DB	1	MKTTLPEFBAAGVWVGDVWLDRTYWGPTSRISPEAPVPVKNVTIEERPGANVAMNT	60
QY	61	ASIGANARLVGLTGIDDAARLSKSLADVNVKCDPVSPHTPTTKRLVLSRNOQLRLD	120
DB	61	ASIGANARLVGLTGIDDAARLSKSLADVNVKCDPVSPHTPTTKRLVLSRNOQLRLD	120
QY	121	FEFGFEGVDQPLHERINQSLISGLVSDYAKGALASVQOMIQLARKGVVLIDPKG	180
DB	121	FEFGFEGVDQPLHERINQSLISGLVSDYAKGALASVQOMIQLARKGVVLIDPKG	180
QY	181	TDPERYRGATLLTPNISEFAVVGKCTBEIEYERGMKLTADYELSLVTRSEOGMSLL	240
DB	181	TDPERYRGATLLTPNISEFAVVGKCTBEIEYERGMKLTADYELSLVTRSEOGMSLL	240
QY	241	OPKAPLHMPPTQAOBYVDVTGADDTYIVGLAATLAAGNSIEBAKCFPANAAGVVGKLTGT	300
DB	241	OPKAPLHMPPTQAOBYVDVTGADDTYIVGLAATLAAGNSIEBAKCFPANAAGVVGKLTGT	300

Qy 301 STVSPILLENNAVRGADTGFVMTTEBELKLA VAAARKGKGVMTNGVFDILHAGHVSYL 360  
Db 301 STVSPILLENNAVRGADTGFVMTTEBELKLA VAAARKGKGVMTNGVFDILHAGHVSYL 360  
Qy 361 ANARKLDRILVAVNSDASTKRLKGDSPVPVPLEQRMIVLGALEAVDMVVSFEEDTPQRL 420  
Db 361 ANARKLDRILVAVNSDASTKRLKGDSPVPVPLEQRMIVLGALEAVDMVVSFEEDTPQRL 420  
Qy 421 IAGILPDLVKGDDYKPEEIKSGKVMANGGEVLVLPFEDCSTTNIKKIQDCKG 477  
Db 421 IAGILPDLVKGDDYKPEEIKSGKVMANGGEVLVLPFEDCSTTNIKKIQDCKG 477

RESULT 2  
US-09-489-9208  
Sequence 9208, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709, 2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 9208  
LENGTH: 499  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-9208

Query Match 93.1%; Score 2227; DB 2; Length 499;  
Best Local Similarity 92.2%; Pred. No. 3.3e-215;  
Matches 437; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MKVTLPFERAGVWVVDVMDLDRYWGPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 60  
Db 23 MKVTLPFERAGVWVVDVMDLDRYWGPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 82  
Qy 61 ASLGANARLVGLTGIDDAARLSSKSLADVNVKCDPVSPVPTITTKLRVLSRNOQLRLD 120  
Db 83 ASLGANARLVGLTGIDDAARLSSKSLADVNVKCDPVSPVPTITTKLRVLSRNOQLRLD 142  
Qy 121 FEEGSEGVDPQLHERINQALSSIGALVLSYKAGALASVQOMQLARKAGVPLIDPKG 180  
Db 143 FEEGSEGVDPQLHERINQALSSIGALVLSYKAGALASVQOMQLARKAGVPLIDPKG 202  
Qy 181 TDFERYRGATLLTPNLSEFEAVVKGCKTEBEIYERGMKLIADYELSLVLRSEQMSLL 240  
Db 203 TDFERYRGATLLTPNLSEFEAVVKGCKTEBEIYERGMKLIADYELSLVLRSEQMSLL 262  
Qy 241 QPGKAPLHPTQAOEYVDVTGAGDTVIGVLAATLAAGNSLEACFPANAAAGVVGKLG 300  
Db 263 QPGKAPLHPTQAOEYVDVTGAGDTVIGVLAATLAAGNSLEACFPANAAAGVVGKLG 322  
Qy 301 STVSPILLENNAVRGADTGFVMTTEBELKLA VAAARKGKGVMTNGVFDILHAGHVSYL 360  
Db 323 STVSPILLENNAVRGADTGFVMTTEBELKLA VAAARKGKGVMTNGVFDILHAGHVSYL 382  
Qy 361 ANARKLDRILVAVNSDASTKRLKGDSPVPVPLEQRMIVLGALEAVDMVVSFEEDTPQRL 420  
Db 383 ANARKLDRILVAVNSDASTKRLKGDSPVPVPLEQRMIVLGALEAVDMVVSFEEDTPQRL 442  
Qy 421 IAGILPDLVKGDDYKPEEIKSGKVMANGGEVLVLPFEDCSTTNIKKIQD 474  
Db 443 IAGILPDLVKGDDYKPEEIKSGKVMANGGEVLVLPFEDCSTTNIKKIQD 496

RESULT 3  
US-09-543-6634  
Sequence 6634, Application US/09543661A

Patent No. 6605709  
GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709, 1002-001  
CURRENT APPLICATION NUMBER: US/09/543,661A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 6634  
LENGTH: 491  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-6634

Query Match 83.9%; Score 2007; DB 2; Length 491;  
Best Local Similarity 82.2%; Pred. No. 4.6e-193;  
Matches 389; Conservative 43; Mismatches 41; Indels 0; Gaps 0;

Qy 1 MKVTLPFERAGVWVVDVMDLDRYWGPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 60  
Db 18 MKVTLPFERAGVWVVDVMDLDRYWGPTSRISPEAPVPVVKVNTIEERPGGANVAMNI 77  
Qy 61 ASLGANARLVGLTGIDDAARLSSKSLADVNVKCDPVSPVPTITTKLRVLSRNOQLRLD 120  
Db 78 ASLGANARLVGLTGIDDAARLSSKSLADVNVKCDPVSPVPTITTKLRVLSRNOQLRLD 137  
Qy 121 FEEGSEGVDPQLHERINQALSSIGALVLSYKAGALASVQOMQLARKAGVPLIDPKG 180  
Db 138 FEEGSEGVDPQLHERINQALSSIGALVLSYKAGALASVQOMQLARKAGVPLIDPKG 197  
Qy 181 TDFERYRGATLLTPNLSEFEAVVKGCKTEBEIYERGMKLIADYELSLVLRSEQMSLL 240  
Db 198 TDFERYRGATLLTPNLSEFEAVVKGCKTEBEIYERGMKLIADYELSLVLRSEQMSLL 257  
Qy 241 QPGKAPLHPTQAOEYVDVTGAGDTVIGVLAATLAAGNSLEACFPANAAAGVVGKLG 300  
Db 258 QPGKAPLHPTQAOEYVDVTGAGDTVIGVLAATLAAGNSLEACFPANAAAGVVGKLG 317  
Qy 301 STVSPILLENNAVRGADTGFVMTTEBELKLA VAAARKGKGVMTNGVFDILHAGHVSYL 360  
Db 318 STVSPILLENNAVRGADTGFVMTTEBELKLA VAAARKGKGVMTNGVFDILHAGHVSYL 377  
Qy 361 ANARKLDRILVAVNSDASTKRLKGDSPVPVPLEQRMIVLGALEAVDMVVSFEEDTPQRL 420  
Db 378 ANARKLDRILVAVNSDASTKRLKGDSPVPVPLEQRMIVLGALEAVDMVVSFEEDTPQRL 437  
Qy 421 IAGILPDLVKGDDYKPEEIKSGKVMANGGEVLVLPFEDCSTTNIKKIQ 473  
Db 438 IAGILPDLVKGDDYKPEEIKSGKVMANGGEVLVLPFEDCSTTNIKKIQ 490

RESULT 4  
US-09-252-991A-32984  
Sequence 32984, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196, 136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 32984  
LENGTH: 526  
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32984

Query Match 56.5%; Score 1351.5; DB 2; Length 526;  
Best Local Similarity 57.9%; Pred. No. 5,46-127;  
Matches 274; Conservative 70; Mismatches 128; Indels 1; Gaps 1;

QY 1 MKTLPFEFAGVWVVDVMDLDRYVGPTRISPEAVPVVKNKTIERRPGAAVNMNT 60  
DB 54 MKLSMPFDDAPVAVVDVMDLDRYVGMGATSRISPEAVPVVQHQHDEPGAAVNMNT 113  
QY 61 ASIGANARLVGLTGIDDAABALSKSLADVNVKCFVSPVTPHTTKLRVLSRNQOLRLD 120  
DB 114 AALGAQALLVGVTRDADSLANSLAAGVADARFGQIDSPITVLRVSRHQOLRLVD 173  
QY 121 FEEGFEVDPOPLHERINQALSSIGALVLSDYAKGALAVQOMIQLARKAGVPLIDPKG 180  
DB 174 FEEFPR-TDAALAVDVESLLAKVVLVLSYKGLQNHQVLIQARARNIPVLADPKG 232  
QY 181 TDERRRGATLLTPNLSEFPAVVKCKTEEEIVERGKCLADYELSLVLTTRSEQKSL 240  
DB 233 KDAVIRGASLITPNLSEFPIVGRCADEALVAKGALMSELDGLVLTGRGKMTLL 292  
QY 241 OPGKAPLHPTQAEVYDTGAGPTVIGVLAATLAAGNSLEBACFPANAAAGVVGKGT 300  
DB 293 RDGPALHLPARAREVDVTGAGDTVSTLAAALAGBELPSAVGLANLAVGKLTGT 352  
QY 301 STVSPLEENAVRGADTGFVMTSEELKLAVAARRKGEKVMTNGVPDILAHGVSYL 360  
DB 353 AALSAPELRABVQEGSESRGVGLDLEQLLALIEDARAHGEKIVFTNGCPDILAHGVSYL 412  
QY 361 ANARKGDRILVAVNSDASTRLKGSRPVNPLEQRITVGALEAVDVMVSPFEDTQRL 420  
DB 413 EOAAQODRLIVGNDASVTRKGVSRPINSVDRRAVLAGLADVMMVSPFEDTQRL 472  
QY 421 IAGLIPDLVKGSDYKPEBIAGSKGVANGGEVVLVNFEDCSTNTIITKIQ 473  
DB 473 LEOVRPDLVKGSDYGEVQVGAQIVKAYGGEVAVLVLENSITTAIVEKIQ 525

## RESULT 5

US-09-477-962-114  
Sequence 114, Application US/09477962  
Patent No. 6927286  
GENERAL INFORMATION:  
APPLICANT: SHEN, BEN  
APPLICANT: DU, LIANGCHENG  
APPLICANT: SANCHEZ, CESAR  
APPLICANT: CHEN, MEI  
APPLICANT: EDWARDS, DANIEL J.  
TITLE OF INVENTION: BLEDOMYCIN GENE CLUSTER COMPONENTS AND THEIR USES  
FILE REFERENCE: 4077-895820US  
CURRENT APPLICATION NUMBER: US/09/477,962  
CURRENT FILING DATE: 2000-01-05  
PRIOR APPLICATION NUMBER: 60/115,435  
PRIOR FILING DATE: 1999-01-06  
PRIOR APPLICATION NUMBER: 60/118,848  
PRIOR FILING DATE: 1999-02-05  
NUMBER OF SEQ ID NOS: 133  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 114  
LENGTH: 498  
TYPE: PRT  
ORGANISM: Streptomyces verticillius  
FEATURE:  
OTHER INFORMATION: ORP9  
US-09-477-962-114

Query Match 25.5%; Score 610.5; DB 2; Length 498;  
Best Local Similarity 37.7%; Pred. No. 2,1e-52;  
Matches 182; Conservative 68; Mismatches 196; Indels 37; Gaps 15;  
QY 13 VMVGVMDLBRVWYGPTSRISPEAVPVVKNKTIERRPGAAVNMNTASIGANARLVGL 72

DB 17 LVAVGDIITDVTWAGTSGLCRSFPVPAVTLTISVAGCGGANVAVNLAALGSEPTLGA 76  
QY 73 TGIDDAABALSKSL-ADVNVKCFVSPVTPHTTKLRVLSRNQOLRLDFEEGFEVD 130  
DB 77 TGDDBRGRRLRALRARDVDTGGLFVQ-PORTVTGRVWADQOMLRLD--EG--GEHP 131

QY 131 QP-----LHERINQALSSIGALVLSDYAKGALAVQOMIQLA--RRAGVVLT--DB 178  
DB 132 LPATDTGSRLLERRAGLLPADAVALVSDYGV-VNEPDVAAALHREIGPSTLVDSR 190  
QY 179 KGTDFRYGATLLTPNLSEF-----BAVVKCKTEEEIVERGKCLADYELSLVLT 231  
DB 191 RPARFTALR-ASAVKNHNAALRLDAGERPPPPARA-DMAALGRRLLRTGAERVALT 248  
QY 232 RSSQKSLQPGKAPLHPTQAE--VYDTGAGDTVIGVLAATLAAGNSLEBACFPANA 289  
DB 249 LDADGSLTFERDRPVTPRTPARAPVTAAGADFTAAALTLAAGADSAAVAELASA 308  
QY 290 AAGVVGKLTSTVSPLEENAVRGADTGFVMTSE-ELKLAVAARRKGEKVMTNGV 348  
DB 309 AAGTAAVAPGTSTWMADELRLRLG--TGKVCRTGTLPARLIDPARDR--RVFTNGC 363  
QY 349 FDLIAGHVSYLANARKGDRILVAVNSDASTRLKGSRPVNPLEQRITVGALEAVDVM 408  
DB 364 FDLIAGHVSCLSRAGELDLVAVNSDASVRLKGPSPRPVPLAERRVLAALSCVDL 423  
QY 409 VVSFEDTQRLIAGLIPDLVKGSDYKPEBIAGSKGVANGGEVVLVNFEDCSTNTI 468  
DB 424 VVFPDDSPALLHEALRPVYAKGDTYTLTBEAPLVQRLGGVHLPSVADTSTTDII 483  
QY 469 KKI 471  
DB 484 RRI 486

## RESULT 6

US-09-489-039A-9668  
Sequence 9668, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 9668  
LENGTH: 342  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9668

Query Match 8.2%; Score 197; DB 2; Length 342;  
Best Local Similarity 26.9%; Pred. No. 5,1e-11;  
Matches 90; Conservative 52; Mismatches 162; Indels 30; Gaps 14;

QY 2 KVTLPFEFAGV--WVGDV-MLDYRYGPTSRISPEAVP--VKNKTIERRPGAA 54  
DB 17 RAALP---RAGIERMMSGKVCVGFSGFNFDMYAKVD-RFPVPEBSLVAGSMTSAGKGA 72  
QY 55 NVAMNTASIGANARLVGLTGIDDAABALSKSLADVNVKCFVSPVTPH-PTTKLRVLSRV 113  
DB 73 NQATLAKGAVVHYIGKIGNDFGHPARHHLKGVGFNNAVTLVABETPGNALIVAGN 132  
QY 114 --QQLRLDPESEFVDPOPLHERINQALSSIG-ALVSDYAKGALAVQOMIQLARKA 170  
DB 133 DARNMLAVD---PGANNVTVDDEIAGCIPALGADAVLVQLENNISAEVQIDAGKA 187  
QY 171 GVPVLIDP---KGTDFRYGATLLTPNLSEFPAVVKCKTEEEIVERGKCLADYELSA 227

Db 188 GATVILNPAQWPEHALLRKVDLTTPNATEAGMTGRRVDSLTAALAAADVLHAQAGRN 247  
 Qy 228 LVTNRSEGMGLTPG-KAPLHPTQAOEVVDYAGDTVIGVLAATLAAGNSLEECF 286  
 Db 248 VITLGSAGALSHEGVSPI--PCPSHPDPTTGAGFAGALAAALACGEPLOAAARF 305  
 Qy 287 ANAAGVVGGLGTSTVSPLEL---ENAVRGRAD 317  
 Db 306 AAAYAAVSVEKQAGASSL-PEVLEAQRLLRAAD 338

## RESULT 7

US-09-252-991A-31882  
 ; Sequence 31882, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252.991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 31882  
 ; LENGTH: 503  
 ; TYPE: PR1  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-31882

Query Match 8.2%; Score 196.5; DB 2; Length 503;  
 Best Local Similarity 26.4%; Pred. No. 1.1e-10;

Matches 88; Conservative 47; Mismatches 163; Indels 35; Gaps 11;

Qy 6 PEFER-----AGVVVGVDMIDRYWYPTSRISPEAPVPVKNVTEERPGG-AAN 55  
 Db 184 PSLERGNKKDAMTKVILVDSLSMDLVVRAPLRPGE---TLAQSFTTIRGKGAN 239  
 Qy 56 VAMNIASIGANARLVLTGIDDAARALSKSLADVNVKDFVS-VPHETTKRLVSRNQ 114  
 Db 240 QAVAAARLGAIVAMIGCLGDAYGQQLYRALQAGIDQGERVAGSSGVALIVVDSS 299  
 Qy 115 QLILDFEPEGEGVDPPFL--HERINQALSSIGALVSDYAKGALASVQOMIQLARKAGV 172  
 Db 300 QMAIVTIVAGGNGHLSPAVLARHEHLLEQ-----AQVWCQLESPLFETVGHVLRRAHALGK 354  
 Qy 173 PVLIDP-----KGTDFERYRGATLLTPNLSEFAVVGKCTEETIVE--RGMKLIADYEL 225  
 Db 355 TVIINPAPATRDVPADMPLVDYLVFNTESEL---CRLVDSLESAGRAAEKLRMG 411  
 Qy 226 SALLVTSEQMSLLQPGKAPLHPTQAOEVVDYAGDTVIGVLAATLAAGNSLEECF 285  
 Db 412 GRVIVTIGAGGALIVGEGRAVE-HFVVARVKALDTTAAGDTFVGGPAAALAGLEAAAIR 470  
 Qy 286 PANAAGVVGKLTGSTVSP--ELENAVGRGA 316  
 Db 471 FGQAALISVTRLAGQTSIPREVERALGEEA 503

## RESULT 8

US-09-134-000C-4524  
 ; Sequence 4524, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134.000C

; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4524  
 ; LENGTH: 529  
 ; TYPE: PR1  
 ; ORGANISM: Enterococcus faecalis  
 US-09-134-000C-4524

Query Match 8.0%; Score 190.5; DB 2; Length 529;  
 Best Local Similarity 24.5%; Pred. No. 4.8e-10;

Matches 81; Conservative 53; Mismatches 126; Indels 71; Gaps 11;

Qy 13 VNVVGVDMIDRYWY-----GPTSRISPEAPVPVKNVTEERPGGAANVAMNIALGA 65  
 Db 219 MIVVGSLSNVNLYSTNLPNGKTNPLSYAKFP-----GKGLNQAAGLTYLGH 268  
 Qy 66 NARLVLTGIDDAARALSKSLADVNVKDFVSVPHTTKRLVSRNQILRLD-----120  
 Db 269 QATLIGLSDPTDANLYKELKRYHTTD-----GTRIDTENGQAYIVETSGDS 320  
 Qy 121 ----FEEGEGVDPPPLHERINQALSSIGALVSDYAKGALASVQOMIQLARKAGVVL 176  
 Db 321 MISILPGANTALTPKIAQCKLFMDASFCLQTEI---PLSAVERKCEIASHGVPIIL 377  
 Qy 177 DP-----KGTDFERYRGATLLTPNLSEFAVVGKCTEES---IVERGMGLIA 221  
 Db 378 KPAAIHIPVNIIEKDF-----FVNEDELLOPRTGLTLEAAVFLMGVAVN- 428  
 Qy 222 DYRISALLVTRSEQMSLLQPGKAPLHPTQAOEVVDYAGDTVIGVLAATLAAGNSLE 281  
 Db 429 -----IVTLGKGV-LKTPQVCHPPTATENIADVSTASDSFISALASYSKGYPT 480  
 Qy 282 EACFPANAAGVVGKLT--STVSPIELN 310  
 Db 481 AATQIAIOAAGFSVSKGVIDSLVDHVTLEN 511

## RESULT 9

US-09-543-681A-7498  
 ; Sequence 7498, Application US/09543681A  
 ; Patent No. 6605709  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GARY BRETTON  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
 ; FILE REFERENCE: 2709.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/543.681A  
 ; CURRENT FILING DATE: 2000-04-05  
 ; PRIOR APPLICATION NUMBER: US 60/128,706  
 ; PRIOR FILING DATE: 1999-04-09  
 ; NUMBER OF SEQ ID NOS: 8344  
 ; SEQ ID NO 7498  
 ; LENGTH: 311  
 ; TYPE: PR1  
 ; ORGANISM: Proteus mirabilis  
 US-09-543-681A-7498

Query Match 7.6%; Score 183; DB 2; Length 311;  
 Best Local Similarity 24.3%; Pred. No. 1.1e-09;

Matches 63; Conservative 47; Mismatches 141; Indels 8; Gaps 5;

Qy 51 GGAANVAMNIASIGANARLVLTGIDDAARALSKSLADVNVKDFVSVPHTTKRLV 110  
 Db 44 GKGAQVAAAGRGANTFLACGNDIGSARAQLTTPDKITDCEFL-IDDEATGVALI 102  
 Qy 111 SRNQQLLR-LDFEPEGVDPPPLHERINQALSSIGALVSDYAKGALASVQOMIQLARK 169  
 Db 103 PAVDQGSNVGTHAGANGRLNREYRYNSVIRKADALLMQ--LESLSLDSVLAQAVAKQ 160  
 Qy 170 AGVPVLIDP--KGTDFERYRGATLLTPNLSEFAVVGKCTEETIVERGMKLIADYELS 226

Db 161 ESQVILINPAPAKLPBELLSDYDITPNETETRYLGVAVVIDESQPSNIIHDKGIE 220  
 Qy 227 ALVTRESEQMSLLOPKGAPLHPETOAGVYDVTAGDVTYIGVLAATLAAGNSLEACFP 286  
 Db 221 TVLITGSRVWVSQNNQGTW-VAPFKVTAIDTIAAGDTFGAFTALLBGLPMMKAIKF 279  
 Qy 287 ANAAGVVGKLGSTVSP 305  
 Db 280 AHAALAAVTRAGAQSPVP 298

## RESULT 10

US-09-489-039A-12757  
 ; Sequence 12757, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; PRIOR FILING DATE: 2000-01-27  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 12757  
 ; LENGTH: 356  
 ; TYPE: PRT  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-12757

Query Match 7.1%; Score 171; DB 2; Length 356;  
 Best Local Similarity 24.2%; Pred. No. 2.3e-08;

Matches 80; Conservative 40; Mismatches 152; Indels 58; Gaps 11;  
 Qy 28 PTRISSEAPVPVVKVNTIERPGG-AAVAMNIAISGANARLVGLGIDDAALASKSL 86  
 Db 63 PVVAIPRGGVAFIE-QIRLNPAGTANAGANINAKGIRTAACACGEDEKADFIILASY 120  
 Qy 87 ADVVVKCDPVV-----VPTHTITTLRLVSNHOOILRLDFEEGEGV-DPOPLH 134  
 Db 121 ARIGDLSLORTALKETSATILPIRNGRRPALHCGASDALFVSAEDAVLDCRFLH 180  
 Qy 135 ERINQALSSIGALVLSYANGALASVQOMIQARKAGPV---LIDPKGDFERYGATL 191  
 Db 181 H-----GGTCLTAAHQGOSA-----RLQAAKARGVTTSPDLIAENEETLILRP--- 226  
 Qy 192 LTPULSEF-----BAVVGKCKTEEEIYERGMKLIADYELALLVTRSEQMSLLOPKAP 246  
 Db 227 LRPVDYFMPSLEBAFLSGETOPEALGRFFLAG---VGTCLIKDGENSGMWLIGRDGP 283  
 Qy 247 LHPETOAGVYDVTAGDVTYIGVLAATLAAGNSLEACFPANAAAGVVGKLGSTVSP 306  
 Db 284 QHIAFWVEAVDTTGGCDSDYCGGFIALAGLSYKAAQDVASNAVALVATGMSDGA----- 339  
 Qy 307 ELENAVRGRADTGFVGTSEELKLAVAAR 336  
 Db 340 -----GVVDMEQTQAFMAHR 355

## RESULT 11

US-09-668-262A-14  
 ; Sequence 14, Application US/09668262A  
 ; Patent No. 6596926  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, Carl  
 ; APPLICANT: Ramodu, Omolayo O.  
 ; APPLICANT: Kinney, Tony  
 ; APPLICANT: Rafalski, Antoni  
 ; TITLE OF INVENTION: Phosphatidylcholine Biosynthetic Enzyme  
 ; FILE REFERENCE: BB1403 US NA  
 ; CURRENT APPLICATION NUMBER: US/09/668,262A

; CURRENT FILING DATE: 2000-09-22  
 ; PRIOR APPLICATION NUMBER: 60/155626  
 ; PRIOR FILING DATE: 1999-09-23  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Microsoft Office 95  
 ; SEQ ID NO 14  
 ; LENGTH: 423  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 US-09-668-262A-14

Query Match 7.1%; Score 169; DB 2; Length 423;  
 Best Local Similarity 30.5%; Pred. No. 4.8e-08;  
 Matches 67; Conservative 24; Mismatches 79; Indels 50; Gaps 9;

Qy 274 LAAGNSLEACFPANAAAGVVGKLGSTVVS-----PIELENAVRGRADTGFVMTSEEL 328  
 Db 1 MEAGAGSSSAKLVAAACVIGIV--LGASVVALHLAGPVAILPALRRRRAPRR----- 50  
 Qy 329 KLAVAARRGKGVNTNGVFDLILHAGVSYLANARKDRLVAVNSDASTRLKGDGR 388  
 Db 51 -----RRGRRRPVAVN-DCGFDMMHYGHCNALRQARALDDELIVGVSDHEITRANGP-- 103  
 Qy 389 PVNPLRQRMIVIGALRAVDVWVSFEEDTP-----QRLIAGILPDLVYK 431  
 Db 104 PVTPLHERLIM--VAAYKMHVHVIDAPYAITDEFMNTLFFNRYNIDYIIHGDDPCLLPD 160  
 Qy 432 GGDYKPEELIAGSKRVANGGEVLVLPEDGCTTNIKKI 471  
 Db 161 GTD-----AVYLAACKV-----GRFKQIRKTEGVSTTDIVGM 192

## RESULT 12

US-10-427-442-14  
 ; Sequence 14, Application US/10427442  
 ; Patent No. 6960704  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, Carl  
 ; APPLICANT: Ramodu, Omolayo O.  
 ; APPLICANT: Kinney, Tony  
 ; APPLICANT: Rafalski, Antoni  
 ; TITLE OF INVENTION: Phosphatidylcholine Biosynthetic Enzyme  
 ; FILE REFERENCE: BB1403 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/427,442  
 ; CURRENT FILING DATE: 2003-05-01  
 ; PRIOR APPLICATION NUMBER: 60/155626  
 ; PRIOR FILING DATE: 1999-09-23  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Microsoft Office 95  
 ; SEQ ID NO 14  
 ; LENGTH: 423  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 US-10-427-442-14

Query Match 7.1%; Score 169; DB 2; Length 423;  
 Best Local Similarity 30.5%; Pred. No. 4.8e-08;  
 Matches 67; Conservative 24; Mismatches 79; Indels 50; Gaps 9;

Qy 274 LAAGNSLEACFPANAAAGVVGKLGSTVVS-----PIELENAVRGRADTGFVMTSEEL 328  
 Db 1 MEAGAGSSSAKLVAAACVIGIV--LGASVVALHLAGPVAILPALRRRRAPRR----- 50  
 Qy 329 KLAVAARRGKGVNTNGVFDLILHAGVSYLANARKDRLVAVNSDASTRLKGDGR 388  
 Db 51 -----RRGRRRPVAVN-DCGFDMMHYGHCNALRQARALDDELIVGVSDHEITRANGP-- 103  
 Qy 389 PVNPLRQRMIVIGALRAVDVWVSFEEDTP-----QRLIAGILPDLVYK 431  
 Db 104 PVTPLHERLIM--VAAYKMHVHVIDAPYAITDEFMNTLFFNRYNIDYIIHGDDPCLLPD 160

QY 432 GGDYKPEEISAKSEWANGGEVLVNFEDGCSITNIIRKI 471  
 Db 161 GTD-----AVALAKKV-----GFFKQIKRTGEGVSTIDYGRM 192

# RESULT 13

US-09-489-039A-8800  
 ; Sequence 8800, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; FILE REFERENCE: 2709, 2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489, 039A  
 ; PRIOR FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 8800  
 ; LENGTH: 665  
 ; TYPE: PRT  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-8800

Query Match 7.0%; Score 167.5; DB 2; Length 665;  
 Best Local Similarity 22.5%; Pred. No. 1.5e-07;  
 Matches 120; Conservative 91; Mismatches 203; Indels 119; Gaps 27;

QY 9 ERAGVAVGVDMMDRYWYGTSPRISPEAPVPPVKNVTEERPGGANVAVNATISLGNAR 68  
 Db 34 KRLDVICIGRAVAVDLVAQIGSRLEDAV-----SPAKYIGSSGNVAFGTALQGLSKA 86  
 QY 69 LVGLTGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRLVLS-RNQQLRLDFEE---G 124  
 Db 87 MLAVGSGEHNKRFLETLNAGVUTVEYI-IDKSRLLRLVNLGDKDETPLIYRNCA 145  
 QY 125 FEQVDPOLHERINQALSSIGALVLS---DYAKGALSVQOMQLARKAGV----- 172  
 Db 146 DMALTPDISL---EYLASSPALAVTGTLSHA-NTBAAVKALEVARRHGLRPLALDIDY 201  
 QY 173 -PVL-IDPKGTDERRGATLLTPNLSE---FEAVVGKKTTEEYVERG-----M 217  
 Db 202 RPLVMGLSLDGGTFLBSGPVTSQLOEVLHLFDLVG---TEEBPHIAGSITDITLAL 258  
 QY 218 KLIDVETSLALVTRSEQMSLQP-----GKAPLHMPTOAQEVYDVYTGADTVIGVLA 271  
 Db 259 KAVNNAATKATLVCRGPMGCVVLESDIPDSWDQVPLQOGVAV-EVLNVLAGDAFMGGL 317  
 QY 272 ATLAGNSLEBACFPANAAAGVVGKLTSTVSP--IELENAVGRADTGFVMTTEELK 329  
 Db 318 RGMINDSEWQACRYANACGALVVSRRHCADAMFTKVELDDYLQ-RAESVPRPVDRLN 376  
 QY 330 LAVAAARERKGVKVMVNTNVPILAGHVSILAN-ARKLGD-----RLIVANSNAST 380  
 Db 377 HLAHVTSRRQWPELCIFAFD-----HKKQLADLARETGRDEACTPOLKLLLAABAAA 431  
 QY 381 KRLKDSRPVNPLEQMTVLG-----ALBAVD---VSPFEDTPQRLIAGILP 426  
 Db 432 QEA-----GLDQRSGLIADGTGYGQRLALNATGKGMWRPFLEBSRRLREHGNIG 483  
 QY 427 DLAVKGD-----YKPEEISAKS-----EVLN---ANGGEVLV 455  
 Db 484 SOLI---DMPLEHVVKCLVFYHPDDPALRAEQDALLEVLVQACNKGHEILL 533

# RESULT 14

US-09-134-001C-4521  
 ; Sequence 4521, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134, 001C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 4521  
 ; LENGTH: 315  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-4521

Query Match 6.9%; Score 165.5; DB 2; Length 315;  
 Best Local Similarity 22.6%; Pred. No. 6.6e-08;  
 Matches 77; Conservative 56; Mismatches 142; Indels 65; Gaps 13;

QY 2 KYTLPEFERAGVAVGVDMMDRYWYGTSPRISPEAPVPPVKNVTEERPGG--AAVAVAN 59  
 Db 3 KVEGEVENKKKVVIVISTVNDKFL--NVKRPK-FGELTHINQAKFEGGKGANQATA 58  
 QY 60 IASLGANARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRLVLSNQQLRL 119  
 Db 59 ASRLADDTTFISKVGDGNANFTLE-----DEKKAQIH-----TOYILTS 98  
 QY 120 DFEF---GFEQVDPQ-----LHERINQALS-----SIGALVLSYAKGAL-----ASVQ 161  
 Db 99 ESEETQOAFITVDEAQNITLVYGANMTLSADVMSVDAFIGADPVAQLEVPPEALE 158  
 QY 162 QMTQLARKAGVPLIDP-----KQTDPERRGATLLTPNLSEFEAVVGKCTER 210  
 Db 159 QAKIRAKQKITVVLNAPALIELPKSLBLTD-----IIPNTEAELLTGISINNE 210  
 QY 211 EYERGMKLIADVELSALVTRSEQMSLQPKAPLHMPTOAQEVYDVYTGADTVIGVLA 270  
 Db 211 SDKETAATYFLDGLISAVLITLGEQGYCAVOBYRM-IPACNVKALIDTTAAGDTFIGAF 269  
 QY 271 AATLAAG-NSLEBACFPANAAAGVVGKLTSTVSPLELE 309  
 Db 270 LSELINDSLNLESHIRLANQASSLTIVQRKAQASIFTRKE 309

# RESULT 15

US-09-489-039A-14043  
 ; Sequence 14043, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; FILE REFERENCE: 2709, 2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489, 039A  
 ; PRIOR FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 14043  
 ; LENGTH: 330  
 ; TYPE: PRT  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-14043

Query Match 6.9%; Score 165; DB 2; Length 330;  
 Best Local Similarity 23.7%; Pred. No. 8e-08;  
 Matches 65; Conservative 44; Mismatches 135; Indels 30; Gaps 6;

QY 51 GGAANYAMNATISAGNARLVGLTGIDDAARALSKSLADVNVKCDPVSVPHTPTTKLRLV 110  
 Db 63 KGKANQAVNAGRGADIACTGDDDIGERRIRQLASDKI-----DVAPYRAV 111  
 QY 111 SRNQQLRLDFEGFEGVDPQLHERINQALS-----SIGALVLSYAKGALA 158

```

Db      112 AGEATGVALIFVNA-BEENVIQIHAGANAALSVQVEAEKERIASAQLLMQ--LESPLK 168
Qy      159 SVQOMIOLAKKAGVPLIDP---KGTDFERYRGATLLTPNLSEFEAVVGCKTEEEIVER 215
Db      169 SVLAATAKIAHHHTTTLVLPAPARELPDELLALVDIITPNETEAKLTGIRVSEDEDAK 228
Qy      216 GKKLIADYEELSLVTRSEQMSLLQPKAKPLHMPTOAEVYDVTGAGDTVIGVLAATLA 275
Db      229 AADVLAAKIGIGTWTILGSRGWLASGES-RRIPGRVQALDTIAAGDTFNGALVTALL 287
Qy      276 AGNSLEACFPANAAGVVGKLGSTVSPLE 309
Db      288 EGTALPEAIRPAHAAAIAVTRKGAQPSVPWRTS 321

```

Search completed: March 18, 2006, 06:08:07  
 Job time : 50 secs



**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 19:03:29 , Search time 7122 Seconds  
(without alignments)  
11445.312 Million cell updates/sec

Title: US-10-520-820-28

Perfect score: 1434  
Sequence: 1 atgaagaatcagctgcgcaga.....aacagatataaagaactaa 1434

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmb1:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_rtr:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hlg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1434	100.0	1434 6	BD268493 Gene iden
2	1434	100.0	1434 6	CQ779270 Sequence
3	1434	100.0	1434 6	AR493607 Sequence
4	1434	100.0	1434 6	AX189075 Sequence
5	1434	100.0	1434 6	AX363670 Sequence
6	1434	100.0	25638 1	ECU28379 Sequence
7	1434	100.0	110000 1	U00096.31
8	1387.6	96.8	110000 1	AE005174_39
9	1387.6	96.8	110000 1	AE005174_40
10	1387.6	96.3	300359 1	BA000007_39
11	1381.2	95.9	292906 1	AE016767
12	1373.2	95.9	292906 1	AE016988
13	1373.2	95.8	110000 1	AE005674_31
14	1011.8	70.6	1500 6	AR385308
15	1008.6	70.3	258050 1	AL627278
16	1008.6	70.1	301574 1	AE016844
17	1005.4	70.1	1431 1	AF155126
18	1003.8	70.0	20604 1	AE008847

C 19	1003.8	70.0	110000 1	AE017220_33	Continuation (34 o
C 20	1002.2	69.9	110000 1	CP000026_31	Continuation (32 o
C 21	999	69.7	1476 1	AF163661	AF163662 Salmone1
C 22	969.8	67.6	1464 1	AF163662	AF163662 Salmone1
C 23	880.8	61.4	110000 1	EX936398_40	Continuation (41 o
C 24	879.2	61.3	10641 1	AE013955	AE013955 Yersinia
C 25	879.2	61.3	208050 1	AJ414144	AJ414144 Yersinia
C 26	879.2	61.3	290924 1	AE017138	AE017138 Yersinia
C 27	873.6	60.9	110000 1	EX950851_40	Continuation (41 o
C 28	870.6	60.7	1353 6	AR385378	Continuation (32 o
C 29	784.8	54.7	348498 1	EX571872	EX571872 Photobact
C 30	784.8	54.7	349980 6	AX770910	AX770910 Sequence
C 31	724.2	50.5	1476 6	AR377456	AR377456 Sequence
C 32	668	46.6	10373 1	AE004313	AE004313 Vibrio ch
C 33	645.2	45.0	1828 1	AY127572	AY127572 Actinobac
C 34	635.2	44.3	1465 1	U32828	U32828 Haemophilus
C 35	635.2	44.3	110000 6	BD426631_15	Continuation (16 o
C 36	635.2	44.3	110000 6	AR274513_15	Continuation (16 o
C 37	635.2	44.3	110000 6	AR632719_15	Continuation (16 o
C 38	628.8	43.8	110000 1	BA000037_05	Continuation (6 of
C 39	626.6	43.7	110000 1	BA000031_04	Continuation (5 of
C 40	625.6	43.6	110000 1	CP000057_14	Continuation (15 o
C 41	625.6	43.6	349980 6	CQ873160	CQ873160 Sequence
C 42	624	43.5	301235 1	AE016799	AE016799 Vibrio vu
C 43	622.4	43.4	110000 1	AE016827_12	Continuation (13 o
C 44	622	43.4	347994 1	CR378664	CR378664 Photobact
C 45	617.4	43.1	10747 1	AE006127	AE006127 Pasteurel

#### ALIGNMENTS

RESULT 1  
BD268493  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BD268493 1434 bp DNA linear PAT 17-JUL-2003  
Gene identified as being necessary in growth of Escherichia coli.  
BD268493.1 GI:33078261  
JP 2002535007-A/165.  
unidentified  
unclassified  
unclassified

1 (pages 1 to 1434)  
Zyskind,J., Ohlsen,K.L., Trawick,J., Forsyth,A.R., Froelich,J.M.,  
Carr,G.V., Yamamoto,R.T. and Xu,H.H.  
Gene identified as being necessary in growth of Escherichia coli  
Patent: JP 2002535007-A 165 22-OCT-2002;  
ELITRA PHARMACEUTICALS INC  
OS B. coli  
PN JP 2002535007-A/165  
PD 22-OCT-2002 JP 2000596147  
PF 27-JAN-2000 US 60/117405  
PI JUDITH ZYSKIND, KARI L. OHLSEN, JOHN TRAWICK, ALLYN R. FORSYTH PI  
JAMIE M. FROELICH,  
PI GRANT J. CARR, ROBERT T. YAMAMOTO, HOWARD H. XU

PC C12N15/09, A61K31/7088, A61K48/00, A61P31/04, C07K14/245, C07K16/12, PC  
C12N1/15,  
PC C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/02, C12Q1/58 PC  
G01N33/15, G01N33/50,  
PC G01N33/53, G01N33/566, C12N15/00, C12N5/00  
CC Gene identified as being necessary in growth of Escherichia  
CC Key coli  
FH source  
FT location/Qualifiers  
FT 1. 1434  
/organism='B. coli'.  
location/Qualifiers  
1. 1434  
/organism='unidentified'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32644'

FEATURES  
source  
ORIGIN

Query Match 100.0%; Score 1434; DB 6; Length 1434;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGTAAACGCTGCGAGGTTTGAACGTGACAGAGTGTGTGTGTGTGTGTGTG 60  
 DB 1 ATGAAAGTAAACGCTGCGAGGTTTGAACGTGACAGAGTGTGTGTGTGTGTGTG 60  
 QY 61 CTGATGCTTACTGTACGCGCCCAACAGTGTATCTGCGCGAAAGCGCGGTGCCGTG 120  
 DB 61 CTGATGCTTACTGTACGCGCCCAACAGTGTATCTGCGCGAAAGCGCGGTGCCGTG 120  
 QY 121 GTTAAAGTAAATACATCGAAGAACGTCCGGGCGCGCGCTAAAGTGGCGATATATC 180  
 DB 121 GTTAAAGTAAATACATCGAAGAACGTCCGGGCGCGCGCTAAAGTGGCGATATATC 180  
 QY 181 GCTTCTCTCGGTCTAAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 DB 181 GCTTCTCTCGGTCTAAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 241 GCGGTGATTAATCTTGTGCGCGACGTCAAGTCAAAATGCACTTCTGTTCTGTACG 300  
 DB 241 GCGGTGATTAATCTTGTGCGCGACGTCAAGTCAAAATGCACTTCTGTTCTGTACG 300  
 QY 301 CATCGACATTTACCAATTAACGGGTACTTTCCGCAACCAAGCTGATCCGTCTGAT 360  
 DB 301 CATCGACATTTACCAATTAACGGGTACTTTCCGCAACCAAGCTGATCCGTCTGAT 360  
 QY 361 TTTGAAGAGGTTTGAAGGTTTGAATCCGAGCGCTGACGAGCGATTAATCAGCG 420  
 DB 361 TTTGAAGAGGTTTGAAGGTTTGAATCCGAGCGCTGACGAGCGATTAATCAGCG 420  
 QY 421 CTGAGTTCGATTTGCGCGCGCTGCTGCTTCTGACCTAAAGGTTGGCGTGAAGGTA 480  
 DB 421 CTGAGTTCGATTTGCGCGCGCTGCTGCTTCTGACCTAAAGGTTGGCGTGAAGGTA 480  
 QY 481 CAGAGATGATCAACTGCGCGCTGAAAGCGGCTTCCGCTGCTGATTCGAAAGGT 540  
 DB 481 CAGAGATGATCAACTGCGCGCTGAAAGCGGCTTCCGCTGCTGATTCGAAAGGT 540  
 QY 541 ACCGATTTTGAAGCGCTACCGCGCGCTACCGCTGTTAAACCGCGAATCTCCGAAATTTGAA 600  
 DB 541 ACCGATTTTGAAGCGCTACCGCGCGCTACCGCTGTTAAACCGCGAATCTCCGAAATTTGAA 600  
 QY 601 GCTGTTGCTGCTTAATGTAAGACCGAAGAAAGATGTTGAGCGCGCATGAACTGAT 660  
 DB 601 GCTGTTGCTGCTTAATGTAAGACCGAAGAAAGATGTTGAGCGCGCATGAACTGAT 660  
 QY 661 GCCGATTAACGAACTCTGCGCTCTGTTAGTGAACCGCTTCCGAAACAGGTTATGCTGCTG 720  
 DB 661 GCCGATTAACGAACTCTGCGCTCTGTTAGTGAACCGCTTCCGAAACAGGTTATGCTGCTG 720  
 QY 721 CAACCGGATTAAGCGCGCTGCTGCTATGCAACCGAAGCGAAGTGTATGACGTTAAC 780  
 DB 721 CAACCGGATTAAGCGCGCTGCTGCTATGCAACCGAAGCGAAGTGTATGACGTTAAC 780  
 QY 781 GGTGCGGCGCAACAGGTAATGCGCTCTGCGCGCAACCGCTGCAAGCGGTAATTCGCTG 840  
 DB 781 GGTGCGGCGCAACAGGTAATGCGCTCTGCGCGCAACCGCTGCAAGCGGTAATTCGCTG 840  
 QY 841 GAAAGAGCTGCTTCTTTGCAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 DB 841 GAAAGAGCTGCTTCTTTGCAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 901 TCCACGCTTTGCGCGATGAGCTGGAATGCTGTACGTGACGTGACATACAGGCTTT 960  
 DB 901 TCCACGCTTTGCGCGATGAGCTGGAATGCTGTACGTGACGTGACATACAGGCTTT 960  
 QY 961 GCGCTGATTAACGGAAGGAACTGAAGTGGCGTGAAGGCGAAGCGCGTAATCGTGTGAA 1020  
 DB 961 GCGCTGATTAACGGAAGGAACTGAAGTGGCGTGAAGGCGAAGCGCGTAATCGTGTGAA 1020

QY 1021 AAGGTGATGATCAACAGGCTGCTTTGACATCTGTCACGCGCGGACGCTCTTATCTG 1080  
 DB 1021 AAGGTGATGATCAACAGGCTGCTTTGACATCTGTCACGCGCGGACGCTCTTATCTG 1080  
 QY 1081 GCAATGCGCGCAAGCTGGGTGACCGCTGATTTGTGCGGTCAACAGGATGCTTCAAC 1140  
 DB 1081 GCAATGCGCGCAAGCTGGGTGACCGCTGATTTGTGCGGTCAACAGGATGCTTCAAC 1140  
 QY 1141 AAGCGCTGAAAGGGAATCCCGCGCGGTAAACCACTCGAACAGGATATGTTGCTG 1200  
 DB 1141 AAGCGCTGAAAGGGAATCCCGCGCGGTAAACCACTCGAACAGGATATGTTGCTG 1200  
 QY 1201 GCGCACTGGAACCGGTGCACTGAGTGTCTGTTGAAGAGCAAGCGGACGCTTG 1260  
 DB 1201 GCGCACTGGAACCGGTGCACTGAGTGTCTGTTGAAGAGCAAGCGGACGCTTG 1260  
 QY 1261 ATGCGCGGATCTTGGCAATCTGCTGTGTGAAGGCGGCGACATTAACCAAGAGATT 1320  
 DB 1261 ATGCGCGGATCTTGGCAATCTGCTGTGTGAAGGCGGCGACATTAACCAAGAGATT 1320  
 QY 1321 GCGGAGTAAAGATCTGCGGCAACGCTGCGGAAAGTGTGCTCAACTTTGAAGAC 1380  
 DB 1321 GCGGAGTAAAGATCTGCGGCAACGCTGCGGAAAGTGTGCTCAACTTTGAAGAC 1380  
 QY 1381 GATTGCTGACGACCAACATCATCAAGAGATCAACAGGATTAAGGCTTAA 1434  
 DB 1381 GATTGCTGACGACCAACATCATCAAGAGATCAACAGGATTAAGGCTTAA 1434

RESULT 2  
 CQ779270 1434 bp DNA linear PAT 11-MAR-2004  
 LOCUS CQ779270  
 DEFINITION Sequence 28 from Patent WO2004005535.  
 ACCESSION CQ779270  
 VERSION CQ779270.1 GI:45381767  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Escherichia coli  
 Bacteri: Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 source  
 Location/Qualifiers  
 1.1434  
 /organism="Escherichia coli"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:562"

ORIGIN  
 Query Match 100.0%; Score 1434; DB 6; Length 1434;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGTAAACGCTGCGAGGTTTGAACGTGACAGAGTGTGTGTGTGTGTGTG 60  
 DB 1 ATGAAAGTAAACGCTGCGAGGTTTGAACGTGACAGAGTGTGTGTGTGTGTGTG 60  
 QY 61 CTGATGCTTACTGTACGCGCCCAACAGTGTATCTGCGCGAAAGCGCGGTGCCGTG 120  
 DB 61 CTGATGCTTACTGTACGCGCCCAACAGTGTATCTGCGCGAAAGCGCGGTGCCGTG 120  
 QY 121 GTTAAAGTAAATACATCGAAGAACGTCCGGGCGCGCTAAAGTGGCGATATATC 180  
 DB 121 GTTAAAGTAAATACATCGAAGAACGTCCGGGCGCGCTAAAGTGGCGATATATC 180  
 QY 181 GCTTCTCTCGGTCTAAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 DB 181 GCTTCTCTCGGTCTAAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

Oy 241 GCGGTAGTAAATCTGCGCCGACGTCGAAAGCTCAATGCGACTTCGTTCTGTAACGACG 300  
 Db 241 GCGGTAGTAAATCTGCGCCGACGTCGAAAGCTCAATGCGACTTCGTTCTGTAACGACG 300  
 Oy 301 CATCCGACCATTAACCAATTACGCGGTACTTCCGCGCAACCAAGCTGATCCGTCTGAT 360  
 Db 301 CATCCGACCATTAACCAATTACGCGGTACTTCCGCGCAACCAAGCTGATCCGTCTGAT 360  
 Oy 361 TTGGAAGAGGTTTCGAAAGGTGTTGATCCGACGCCGCTGACGAGCGGATTAATCAGCG 420  
 Db 361 TTGGAAGAGGTTTCGAAAGGTGTTGATCCGACGCCGCTGACGAGCGGATTAATCAGCG 420  
 Oy 421 CTGAGTTGATGCGCGCTGCTGCTTCTGATACGCGCAAGGTCGCGCGCAAGCGTA 480  
 Db 421 CTGAGTTGATGCGCGCTGCTGCTTCTGATACGCGCAAGGTCGCGCGCAAGCGTA 480  
 Oy 481 CAGCAGATGATCCAACTGCGCGCTAAAGCGGCTGTTCCGCTGCTGATTTGATCCAAAAGT 540  
 Db 481 CAGCAGATGATCCAACTGCGCGCTAAAGCGGCTGTTCCGCTGCTGATTTGATCCAAAAGT 540  
 Oy 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAAGCGGATCTCTCGAATTTGAA 600  
 Db 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAAGCGGATCTCTCGAATTTGAA 600  
 Oy 601 GCTGTGTGCTGTAATGTAAGACCGAAGAGAGATGTTGAGCGCGGCAATGAACTGATT 660  
 Db 601 GCTGTGTGCTGTAATGTAAGACCGAAGAGAGATGTTGAGCGCGGCAATGAACTGATT 660  
 Oy 661 GCCGATTAACGAATCTCGCGCTCTGTTAGTGAACCGCTTCGAAACAGGATATGCTGCTG 720  
 Db 661 GCCGATTAACGAATCTCGCGCTCTGTTAGTGAACCGCTTCGAAACAGGATATGCTGCTG 720  
 Oy 721 CAACCGGTAAGCGCGCTGCTGATATGCAACCAAGCGCAAGAGTATGACGTTAAC 780  
 Db 721 CAACCGGTAAGCGCGCTGCTGATATGCAACCAAGCGCAAGAGTATGACGTTAAC 780  
 Oy 781 GGTGCGGCGCAACAGGTAATGCGCTGCTGCGCAACCGCTGCAAGCGGTAATTCCTG 840  
 Db 781 GGTGCGGCGCAACAGGTAATGCGCTGCTGCGCAACCGCTGCAAGCGGTAATTCCTG 840  
 Oy 841 GAGAAACCTGCTTCTTTGCAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 Db 841 GAGAAACCTGCTTCTTTGCAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 Oy 901 TCACGCGTTTCGCGGATCGAGCTGAAATGCTGTAAGTGAAGTGAAGTGAAGTGAAGT 960  
 Db 901 TCACGCGTTTCGCGGATCGAGCTGAAATGCTGTAAGTGAAGTGAAGTGAAGTGAAGT 960  
 Oy 961 GCGCTGTAAGCGGAGAGGAACTGAAGTGGCGCTGAGCGGCGCTGAACCTGCTGTA 1020  
 Db 961 GCGCTGTAAGCGGAGAGGAACTGAAGTGGCGCTGAGCGGCGCTGAACCTGCTGTA 1020  
 Oy 1021 AAGTGTGATGACCAACGCTGCTTGAATCTGCAAGCGCGGCAAGCTCTTATCTG 1080  
 Db 1021 AAGTGTGATGACCAACGCTGCTTGAATCTGCAAGCGCGGCAAGCTCTTATCTG 1080  
 Oy 1081 GCAAAATCCCGCAAGCTGAGTGAACCTGTAATGCTGCAACAGCGATGCTTCAAC 1140  
 Db 1081 GCAAAATCCCGCAAGCTGAGTGAACCTGTAATGCTGCAACAGCGATGCTTCAAC 1140  
 Oy 1141 AAGCGCTGAAGAGGAGTTCGCGCGCGCTGAACCACTGCAACAGCGTATGATGCTG 1200  
 Db 1141 AAGCGCTGAAGAGGAGTTCGCGCGCGCTGAACCACTGCAACAGCGTATGATGCTG 1200  
 Oy 1201 GCGCACTGGAAGCGCTGCACTGCGGTAGTCTGTTGAAGAGCAACGCGCGACGCTG 1260  
 Db 1201 GCGCACTGGAAGCGCTGCACTGCGGTAGTCTGTTGAAGAGCAACGCGCGACGCTG 1260  
 Oy 1261 ATGCGCGGATCTTGGCAAGATCTGCTGTTGAAGAGCGCGCACTATTAACCAAGAGAT 1320  
 Db 1261 ATGCGCGGATCTTGGCAAGATCTGCTGTTGAAGAGCGCGCACTATTAACCAAGAGAT 1320

Oy 1321 GCGGAGTAAAGTCTGCGCGCAACGCTGCGGAGTGTGCTGCACTTTGAAGAC 1380  
 Db 1321 GCGGAGTAAAGTCTGCGCGCAACGCTGCGGAGTGTGCTGCACTTTGAAGAC 1380  
 Oy 1381 GGTGCTGACGACCAACATCATCAAGAGATCAACAGGATTAAGGCTAA 1434  
 Db 1381 GGTGCTGACGACCAACATCATCAAGAGATCAACAGGATTAAGGCTAA 1434  
 RESULT 3  
 AR493607  
 LOCUS AR493607 1434 bp DNA linear PAT 15-MAY-2004  
 DEFINITION Sequence 165 from patent US 6720139.  
 ACCESSION AR493607  
 VERSION AR493607.1 GI:47265540  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (pages 1 to 1434)  
 AUTHORS Zytkind, J., Ohlsen, K.L., Trawick, J., Forsyth, R.A., Froelich, J.M.,  
 Carr, G.U., Yamamoto, R.T., and Xu, H.H.  
 TITL Genes identified as required for proliferation in *Escherichia coli*  
 JOURNAL Patent: US 6720139-A 165 13-APR-2004;  
 Birta Pharmaceuticals, Inc.; San Diego, CA  
 FEATURES  
 source 1..1434  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Query Match 100.0%; Score 1434; DB 6; Length 1434;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ATGAAGTAAAGCTGCGCAAGTTTGAACGTGACAGAGTGAAGTGTGCTGATGATG 60  
 Db 1 ATGAAGTAAAGCTGCGCAAGTTTGAACGTGACAGAGTGAAGTGTGCTGATGATG 60  
 Oy 61 CTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 61 CTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Oy 121 GTTAAAGTAAATCAATCGAAGAAAGTCCGCGCGCGCTGAACGCGGATGAATATC 180  
 Db 121 GTTAAAGTAAATCAATCGAAGAAAGTCCGCGCGCGCTGAACGCGGATGAATATC 180  
 Oy 181 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Db 181 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 Oy 241 GCGCTGATTAATCTGCGCGCAAGTCAACGCTCAATGCGACTTCTGTAACGACG 300  
 Db 241 GCGCTGATTAATCTGCGCGCAAGTCAACGCTCAATGCGACTTCTGTAACGACG 300  
 Oy 301 CATCCGACCATTAACCAATTACGCGGTACTTCCGCGCAACCAAGCTGATCCGTCTGAT 360  
 Db 301 CATCCGACCATTAACCAATTACGCGGTACTTCCGCGCAACCAAGCTGATCCGTCTGAT 360  
 Oy 361 TTGGAAGAGGTTTCGAAAGGTGTTGATCCGACGCCGCTGACGAGCGGATTAATCAGCG 420  
 Db 361 TTGGAAGAGGTTTCGAAAGGTGTTGATCCGACGCCGCTGACGAGCGGATTAATCAGCG 420  
 Oy 421 CTGAGTTGATGCGCGCTGCTGCTTCTGATACGCGCAAGGTCGCGCGCAAGCGTA 480  
 Db 421 CTGAGTTGATGCGCGCTGCTGCTTCTGATACGCGCAAGGTCGCGCGCAAGCGTA 480  
 Oy 481 CAGCAGATGATCCAACTGCGCGCTAAAGCGGCTGTTCCGCTGCTGATTTGATCCAAAAGT 540  
 Db 481 CAGCAGATGATCCAACTGCGCGCTAAAGCGGCTGTTCCGCTGCTGATTTGATCCAAAAGT 540  
 Oy 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAAGCGGATCTCTCGAATTTGAA 600

Db 541 ACCGATTTTGAGCCCTACCCGCGCTACGCTGTTAAAGCCGAAATCTCTCGGAATTTGAA 600  
 Qy 601 GCTGTTCGGTAAATATAGACCGAAGAGATTTGTAGGCGGCGATGAATCTGAT 660  
 Db 601 GCTGTTCGGTAAATATAGACCGAAGAGATTTGTAGGCGGCGATGAATCTGAT 660  
 Qy 661 GCCGATTACGAATCTCGGCTCTGTGTAGTGAACCGGTTCCGAACAAGGATGTCTGCTG 720  
 Db 661 GCCGATTACGAATCTCGGCTCTGTGTAGTGAACCGGTTCCGAACAAGGATGTCTGCTG 720  
 Qy 721 CAACCGGGTAAAGCGCGCTGCAATATGCCAACCAAGCGCGAGAGTGTATGACTTACC 780  
 Db 721 CAACCGGGTAAAGCGCGCTGCAATATGCCAACCAAGCGCGAGAGTGTATGACTTACC 780  
 Qy 781 GGTGCGGGCGACAGGATGATGGGCTCTGCGCGCAACCGCTGGAAGGGGATTAATGCTG 840  
 Db 781 GGTGCGGGCGACAGGATGATGGGCTCTGCGCGCAACCGCTGGAAGGGGATTAATGCTG 840  
 Qy 841 GAAAGAGCTGCTCTTCTTGGCAATGCGCGCGCTGCGGCTGCTGCTGCTGCTGCTG 900  
 Db 841 GAAAGAGCTGCTCTTCTTGGCAATGCGCGCGCTGCGGCTGCTGCTGCTGCTGCTG 900  
 Qy 901 TCCACGTTTTCGCGATGAGCTGAAATGCTGTAGTGAAGTGTGATGACGCTTT 960  
 Db 901 TCCACGTTTTCGCGATGAGCTGAAATGCTGTAGTGAAGTGTGATGACGCTTT 960  
 Qy 961 GGGCGATGACCGAAGAGAACTGAAGCTGCGGAGCGGAGCGGCTAAACGATGAA 1020  
 Db 961 GGGCGATGACCGAAGAGAACTGAAGCTGCGGAGCGGAGCGGCTAAACGATGAA 1020  
 Qy 1021 AAGGTGTATGACCAACGCTGTCTTGAACATCTGTGACGCGCGGCAAGTCTTATCTG 1080  
 Db 1021 AAGGTGTATGACCAACGCTGTCTTGAACATCTGTGACGCGCGGCAAGTCTTATCTG 1080  
 Qy 1081 GCAAAATGCGGCAAGCTGAGTGTGATGCTGTGCTGCAACAGGATGCTTCAAC 1140  
 Db 1081 GCAAAATGCGGCAAGCTGAGTGTGATGCTGTGCTGCAACAGGATGCTTCAAC 1140  
 Qy 1141 AAGCGGCTGAAGAGGAAATCCCGCGCGGTAAACCACTGCAACAGGATGATGCTG 1200  
 Db 1141 AAGCGGCTGAAGAGGAAATCCCGCGCGGTAAACCACTGCAACAGGATGATGCTG 1200  
 Qy 1201 GGGCACTGGAAGCGGTGATGCTGTGATGCTGTGAAAGAGACACGCGCGAGCTTG 1260  
 Db 1201 GGGCACTGGAAGCGGTGATGCTGTGATGCTGTGAAAGAGACACGCGCGAGCTTG 1260  
 Qy 1261 ATGCGCGGATCTTGTGCAATCTGTGTGAAAGCGCGCACTTAAACCAAGAGATTT 1320  
 Db 1261 ATGCGCGGATCTTGTGCAATCTGTGTGAAAGCGCGCACTTAAACCAAGAGATTT 1320  
 Qy 1321 GCCGAGAGTAAAGAGTCTGAGCGCAACGCTGCGAAGTGTGTCTCACTTTGAAGAC 1380  
 Db 1321 GCCGAGAGTAAAGAGTCTGAGCGCAACGCTGCGAAGTGTGTGTCTCACTTTGAAGAC 1380  
 Qy 1381 GGTGTCTGACGACCAATCATCAAGAGATCCAAAGATTAAGGCTAA 1434  
 Db 1381 GGTGTCTGACGACCAATCATCAAGAGATCCAAAGATTAAGGCTAA 1434

RESULT 4  
 AX189075 1434 bp DNA linear PAT 08-AUG-2001  
 LOCUS Sequence 276 from Patent WO0148209.  
 DEFINITION AX189075  
 ACCESSION AX189075  
 VERSION AX189075.1 GI:15142798  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Escherichia coli  
 Escherichia coli  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.  
 REFERENCE  
 1 Forsyth, R.A., Ohlsen, K.L. and Zyakind, J.W.  
 Genes identified as required for proliferation of E. coli

JOURNAL Patent: WO 0148209-A 276 05-JUL-2001;  
 FEATURES Biftra Pharmaceuticals, Inc. (US)  
 source Location/Qualifiers  
 1..1434  
 /organism="Escherichia coli"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:562"  
 1..1434  
 CDS  
 /note="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC50756.1"  
 /db\_xref="GI:15142798"  
 translation="MKVTLPEPERAGVWVGVMDLDRVYGPTRISPEAPVAVKVN  
 TIERPQGANVAMNIAISLGAANALVGLTIDDAALSKSLADVAVKCPVSPHP  
 TTKLAVLRNQLIRLDPBEGPEGDPQPIHRIINQALVSLGALVSDKAGLAVS  
 OGMITQLARKAPVILDPKGTDPERVRGATLTLPNISREPAVYKCKTEPEYERGMK  
 LINDYELSLVTRSEQMSLIDPGAPLHPFOQDEVLDVYAGDTVIGLAAITLAA  
 GNSLEBACFPANNAAGVVGKSTVSPLEENAVRGADVTGVTWTEBELLAVALA  
 ARKRGKRVWTVNGVPDILHAGVSYLANNAKLDRLIVAVNSDASTKRLGDSRPNVP  
 LEORMLVGLALEAVDVFVPEBDTPQRLIAGLIPDLVKGQDYKPEIAGSKWMANG  
 GEVLVNFEDGCTTNIKKIQDKKG"

ORIGIN  
 Query Match 100.0%; Score 1434; DB 6; Length 1434;  
 Best Local Similarity 100.0%; Pred.No. 0;  
 Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAGTAAACGCTGCCAAGGTTGAACGTCGAGAGTGAATGGTGTGTGATGATG 60  
 Db 1 ATGAAAGTAAACGCTGCCAAGGTTGAACGTCGAGAGTGAATGGTGTGTGATGATG 60  
 Qy 61 CTGGAATGCTTACTGTGATGCGGCCCAACAGTCTATCTGCGCGGAAGCGCGGTGCCG 120  
 Db 61 CTGGAATGCTTACTGTGATGCGGCCCAACAGTCTATCTGCGCGGAAGCGCGGTGCCG 120  
 Qy 121 GTTAAAGTAAATCAATCGAAGAACTCCGCGCGCGGCTTAAACGTGCGATGATATC 180  
 Db 121 GTTAAAGTAAATCAATCGAAGAACTCCGCGCGCGGCTTAAACGTGCGATGATATC 180  
 Qy 181 GCTTCTCTCGGTCTAATGACGCGCTGTCGGGTGAACGCGGCAATGACATGACGCGCG 240  
 Db 181 GCTTCTCTCGGTCTAATGACGCGCTGTCGGGTGAACGCGGCAATGACATGACGCGCG 240  
 Qy 241 GCGCTGAGTAAATCTCTGCGCGACGTCACAGTCAAAATGCGACTTCTGTATCCGACG 300  
 Db 241 GCGCTGAGTAAATCTCTGCGCGACGTCACAGTCAAAATGCGACTTCTGTATCCGACG 300  
 Qy 301 CATCCGACATTAACCAATTAACGGGTACTTCCGCAACCAACAGCTGATCCGTGAT 360  
 Db 301 CATCCGACATTAACCAATTAACGGGTACTTCCGCAACCAACAGCTGATCCGTGAT 360  
 Qy 361 TTTGAAGAGGTTTCCAAAGGTGTGATCCGCAACCGCTGACAGAGCGATTAATCAGGCG 420  
 Db 361 TTTGAAGAGGTTTCCAAAGGTGTGATCCGCAACCGCTGACAGAGCGATTAATCAGGCG 420  
 Qy 421 CTGAGTTCGATTTGAGCGGCTGTGCTTTCTGACTACCCCAAGTGTGCGCAAGGCTA 480  
 Db 421 CTGAGTTCGATTTGAGCGGCTGTGCTTTCTGACTACCCCAAGTGTGCGCAAGGCTA 480  
 Qy 481 CAGCAGATGATCCAACTGCGCGGTAAAGCGGGTGTTCGGTCTGATGATCCAAAGGT 540  
 Db 481 CAGCAGATGATCCAACTGCGCGGTAAAGCGGGTGTTCGGTCTGATGATCCAAAGGT 540  
 Qy 541 ACCGATTTTGAAGGCTACCGCGCGCTACGCTGTAAACGCAATCTCTGGAATTTGAA 600  
 Db 541 ACCGATTTTGAAGGCTACCGCGCGCTACGCTGTAAACGCAATCTCTGGAATTTGAA 600  
 Qy 601 GCTGTTCGGTAAATATAGACCGAAGAGATTTGTAGGCGGCGATGAATCTGAT 660  
 Db 601 GCTGTTCGGTAAATATAGACCGAAGAGATTTGTAGGCGGCGATGAATCTGAT 660  
 Qy 661 GCCGATTACGAATCTCGGCTCTGTGTAGTGAACCGGTTCCGAACAAGGATGTCTGCTG 720

Db 661 GCCATTACGAACTCTGGCTCTGTTAGTACCCGTTCCGACAGGGTATGCTGCTG 720  
 Qy 721 CAACCGGGTAAAGCCGCTGCATATGCCAACCAAGCGAGAAAGTATGACGTTAAC 780  
 Db 721 CAACCGGGTAAAGCCGCTGCATATGCCAACCAAGCGAGAAAGTATGACGTTAAC 780  
 Qy 781 GGTGGGGGCGACAGGTGATTTGGGCTCTGGCGGCAACGCTGGCAGCGGGTAAATTCGCTG 840  
 Db 781 GGTGGGGGCGACAGGTGATTTGGGCTCTGGCGGCAACGCTGGCAGCGGGTAAATTCGCTG 840  
 Qy 841 GAAGAACCTGCTCTTTGCAATGCGGCGGCTGAGTGTGTGCGGCAATGGGAAC 900  
 Db 841 GAAGAACCTGCTCTTTGCAATGCGGCGGCTGAGTGTGTGCGGCAATGGGAAC 900  
 Qy 901 TCACAGGTTTCGCGGATCGAGCTGGAATACTGTACGTGACGTGCAGATACAGGCTTT 960  
 Db 901 TCACAGGTTTCGCGGATCGAGCTGGAATACTGTACGTGACGTGCAGATACAGGCTTT 960  
 Qy 961 GCGCTGATGACCGAAGAGAACTGAAGTGGCGGTAGGCGCAGCGGCTAAACGTGTGAA 1020  
 Db 961 GCGCTGATGACCGAAGAGAACTGAAGTGGCGGTAGGCGCAGCGGCTAAACGTGTGAA 1020  
 Qy 1021 AAAGTGTGATGACCAACGCTGTCTTTGACATCTGACAGCGGCGGCACTCTTTATCTG 1080  
 Db 1021 AAAGTGTGATGACCAACGCTGTCTTTGACATCTGACAGCGGCGGCACTCTTTATCTG 1080  
 Qy 1081 GCAAAATGCGCGCAAGCTGGGTGACCGCTTGAATTTGTCGCTCAACAGGATGCTCCACC 1140  
 Db 1081 GCAAAATGCGCGCAAGCTGGGTGACCGCTTGAATTTGTCGCTCAACAGGATGCTCCACC 1140  
 Qy 1141 AAAGGCTGAAAGGGGATTTCCCGCCCGGTAAACCACTGCAACAGCGTATGATGTGCTG 1200  
 Db 1141 AAAGGCTGAAAGGGGATTTCCCGCCCGGTAAACCACTGCAACAGCGTATGATGTGCTG 1200  
 Qy 1201 GCGGCACTGAAAGCGGTGACCTGGGTAGTGTCTTTGAAAGAGACAGCGCGCAGCGCTTG 1260  
 Db 1201 GCGGCACTGAAAGCGGTGACCTGGGTAGTGTCTTTGAAAGAGACAGCGCGCAGCGCTTG 1260  
 Qy 1261 ATGCGCGGAGCTTTGCAAGATCTGTGTGTAAGAGCGGCGCACTAATAACGAGAGATT 1320  
 Db 1261 ATGCGCGGAGCTTTGCAAGATCTGTGTGTAAGAGCGGCGCACTAATAACGAGAGATT 1320  
 Qy 1321 GCGGAGTAAAGAGTCTGGGCCAACGCTGGCCGAAGTGTGTGCTCACTTTGAAGAC 1380  
 Db 1321 GCGGAGTAAAGAGTCTGGGCCAACGCTGGCCGAAGTGTGTGCTCACTTTGAAGAC 1380  
 Qy 1381 GGTGTCTGACGACCAATCATCAAGAGATCCAAAGATTAATAAAGCTAA 1434  
 Db 1381 GGTGTCTGACGACCAATCATCAAGAGATCCAAAGATTAATAAAGCTAA 1434  
 RESULT 5  
 AX363670 1434 bp DNA linear PAT 15-FEB-2002  
 LOCUS AX363670  
 DEFINITION Sequence 165 from Patent EP1178052.  
 ACCESSION AX363670.1 GI:18695784  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Escherichia coli  
 Escherichia coli  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.  
 REFERENCE  
 1  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 1. 1434  
 /organism="Escherichia coli"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:562"

ORIGIN  
 Query Match 100.0%; Score 1434; DB 6; Length 1434;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 1 ATGAAAGTAAAGCGGCGAGGTTTGAACGTGACAGAGTATGGGTGATGATGATG 60  
 1 ATGAAAGTAAAGCGGCGAGGTTTGAACGTGACAGAGTATGGGTGATGATGATG 60  
 Qy 61 CTGATCGTTACTGTGACGCGCCCAACGATCGTATCTGCGGAGCGCGGCTCCGCTG 120  
 Db 61 CTGATCGTTACTGTGACGCGCCCAACGATCGTATCTGCGGAGCGCGGCTCCGCTG 120  
 Qy 121 GTTAAAGTAAATACATGAAAGACGTCCGCGGCGCGGCTTAACTGTGGATGATATC 180  
 Db 121 GTTAAAGTAAATACATGAAAGACGTCCGCGGCGCGGCTTAACTGTGGATGATATC 180  
 Qy 181 GCTTCTGCGGTGCTAATGACGCGCTGCGGGTGAAGGCGCAATTGACATGACGCGC 240  
 Db 181 GCTTCTGCGGTGCTAATGACGCGCTGCGGGTGAAGGCGCAATTGACATGACGCGC 240  
 Qy 241 GCGCTGAGTAAATCTGCGCGCAAGTCAACGTCAATGCGACTTCTGTGACCAAGC 300  
 Db 241 GCGCTGAGTAAATCTGCGCGCAAGTCAACGTCAATGCGACTTCTGTGACCAAGC 300  
 Qy 301 CATCCGACATTAACCAATTAACGAGTATCTTCCGCAACCAAGCTGATCCGTGAT 360  
 Db 301 CATCCGACATTAACCAATTAACGAGTATCTTCCGCAACCAAGCTGATCCGTGAT 360  
 Qy 361 TTTGAAAGGTTTCAAGGTTGATCCGACGCGCTGACAGCGGATTAATCAAGCG 420  
 Db 361 TTTGAAAGGTTTCAAGGTTGATCCGACGCGCTGACAGCGGATTAATCAAGCG 420  
 Qy 421 CTGAGTTCGATTTGAGCGCGCTGCTGCTTCTGACTACCCAAAGTGCCTGCAAGCTA 480  
 Db 421 CTGAGTTCGATTTGAGCGCGCTGCTGCTTCTGACTACCCAAAGTGCCTGCAAGCTA 480  
 Qy 481 CAGCAGATGATCCAACTGGCGCGTAAAGCGGCTGCTCCGCTGATGATCCAAAGGT 540  
 Db 481 CAGCAGATGATCCAACTGGCGCGTAAAGCGGCTGCTCCGCTGATGATCCAAAGGT 540  
 Qy 541 ACCGATTTGAGGCTACGCGCGCGCTAGGCTGTTAACGCGGATCTCGGATTTGAA 600  
 Db 541 ACCGATTTGAGGCTACGCGCGCGCTAGGCTGTTAACGCGGATCTCGGATTTGAA 600  
 Qy 601 GCTGTGTGCGTAAATGTAAGCAAGAGATTTGTAAGCGCGCATGAATCTGAT 660  
 Db 601 GCTGTGTGCGTAAATGTAAGCAAGAGATTTGTAAGCGCGCATGAATCTGAT 660  
 Qy 661 GCGGATTAAGCACTCTGCGCTCTGTTAGTGAACCGGTTCCGAACAGGATATGCTGCTG 720  
 Db 661 GCGGATTAAGCACTCTGCGCTCTGTTAGTGAACCGGTTCCGAACAGGATATGCTGCTG 720  
 Qy 721 CAACCGGTTAAAGCGCGCTGCATATGCCAACCAAGCGAGAAAGTATGACGTTAAC 780  
 Db 721 CAACCGGTTAAAGCGCGCTGCATATGCCAACCAAGCGAGAAAGTATGACGTTAAC 780  
 Qy 781 GGTGGGGGCGACAGGTGATTTGGGCTCTGGCGGCAACGCTGGCAGCGGGTAAATTCGCTG 840  
 Db 781 GGTGGGGGCGACAGGTGATTTGGGCTCTGGCGGCAACGCTGGCAGCGGGTAAATTCGCTG 840  
 Qy 841 GAAGAACCTGCTCTTTGCAATGCGGCGGCTGAGTGTGTGCGGCAATGGGAAC 900  
 Db 841 GAAGAACCTGCTCTTTGCAATGCGGCGGCTGAGTGTGTGCGGCAATGGGAAC 900  
 Qy 901 TCACAGGTTTCGCGGATCGAGCTGGAATACTGTACGTGACGTGCAGATACAGGCTTT 960  
 Db 901 TCACAGGTTTCGCGGATCGAGCTGGAATACTGTACGTGACGTGCAGATACAGGCTTT 960  
 Qy 961 GCGCTGATGACCGAAGAGAACTGAAGTGGCGGTAGGCGCAGCGGCTAAACGTGTGAA 1020  
 Db 961 GCGCTGATGACCGAAGAGAACTGAAGTGGCGGTAGGCGCAGCGGCTAAACGTGTGAA 1020





misc\_feature 10928..15584  
/note="similar to GenBank Accession Number M16194  
(ECORPSU)"  
misc\_feature 10978..15584  
/note="similar to GenBank Accession Number L14781  
(ECOTATDAB)"  
gene 11258..12169  
/gene="tcdA"  
11258..12169  
/gene="tcdA"  
/EC number="4.2.1.32"  
/codon\_start=1  
/transl\_table=1  
/product="tartarate dehydratase"  
/protein\_id="AA089141.1"  
/db\_xref="GI:882584"  
/translation="MMSBNKQAQAVNKLTEIVANFTAMISTMPDDVDLTKQLKDAE  
TSSMGIYHTEMNMOKAIIDLRPAQDGTGEIMFVKSREPLLEQSLIKQAVE  
EATYKAPLRNAAVEIPDEVNTKNTGSGVPMWMDIIPNDDAIEIYVMAAGGCTLP  
RSYVAMPSEGEVGVKVPENISITLANAPVLIYVGIATSVETAIVSRKAIIRPI  
GSRHPNKAABLRLEEGNRLGIGOGILTGNSVWGCHIESAARPSITIGAVSTG  
CMARRGTLVHDLTEENLSHRSAL"  
CDS 12166..12771  
/gene="tcdB"  
12166..12771  
/gene="tcdB"  
/EC number="4.2.1.32"  
/codon\_start=1  
/transl\_table=1  
/product="tartarate dehydratase"  
/protein\_id="AA089142.1"  
/db\_xref="GI:882585"  
/translation="WKKITLTPIKAEPLQDIRGVYVLTGLVTCEDVCHRLIEK  
RPIPYDLNGAIFHAGPIVRKNGDKMEVSGPTTSRMSEFKEEIQGVNLVKEK  
GPIVSIIDTGNNTLAEKKLPAERDPIVEICEHVIYK"  
12819..14282  
/note="alternate name y5jB; ORF\_0487"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AA09143.1"  
/db\_xref="GI:882586"  
/translation="MKPSTMMRYLAPLAVIAIITALLPVPAGLHNTMLYPVFTGYI  
VGLIRPVGAVVAMVAGISITILSPDLNPGKRTAKLSMAVSPSNV  
IWLIFAPMFGTEKTEGLRILNIVKMGHRTPLGTYMPSSEILLPVPSNA  
RGAGIYPIIRNLPPLVQSPNDSSRSISYIMMGIVADVSAIFLAPMAENLL  
IGLMSASHATLSGDFLQMLPLSILVLVPLVAVLPVLSKSDQVPRAPNELL  
OANGPLCSREKMLGLMVALVMIFGDVLIDAMVGVVALLMLLISMPDIYSN  
KAMNVFPMILASITLALGLNNGSIFPMKILAGSISYPTMMVALLIVPDIYSN  
PMSATVYTSALAPMIAALAMPRIPLPCLMGAIGLSILPYATGSPPIYIG  
SGYLPADYVRLGIFGLIFVLVITGLLMPVLL"  
CDS complement (14325..15338)  
/note="ORF\_0337"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AA089144.1"  
/db\_xref="GI:882587"  
/translation="MVLGIETSCDETGAITYDEKGLANQYSQVKLADYGVVP  
ELASRDVIRKTVPLIQALKEGSLADIDAVATAPGVLGALLVATGSRSLAFW  
DVAPILPVHNEGHLLAPLBDNPPSPFVLLVSGHTOLISVTGICQYLLSIDI  
AABAPKTKAKLGLDYRGGPLSKMAAGTAGFRVPRMTBRTGQDPSGLKIPDA  
ANTIRDGTDDQTRADYARAFEDAVYDTLMIKCRALDQDGFRELVAAGSVANRTIR  
AKLAEMMKRRGEVYARPECTDNGMIAVAGNVRKAGATADLGVSVERPMLAEL  
PAA"  
misc\_feature 14862..15917  
/note="similar to GenBank Accession Number J01687  
(ECORPSRPO)"  
gene 15576..15791  
/gene="rpsU"  
15576..15791  
/gene="rpsU"  
/note="CG site No. 211"  
/codon\_start=1

/transl\_table=1  
/product="30S ribosomal subunit protein S21"  
/protein\_id="AA089145.1"  
/db\_xref="GI:882588"  
/translation="MPLIVRENEPPDALRFRKSCERAGVLAVERREFEKPTTE  
RKRKASAVRPAKXKLARENARTRLY"  
gene 15902..17647  
/gene="dnaG"  
15902..17647  
/gene="dnaG"  
/function="DNA biosynthesis; primase"  
/note="CG site No. 847; alternate gen name dnaP, parB"  
/codon\_start=1  
Query Match 100.0%; Score 1434; DB 1; Length 25638;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAGTAAACGCTGCAGAGTTTGAACGTGCAGAGTGAATGGTGGTGAATGATG 60  
DB 1548 ATGAAGTAAACGCTGCAGAGTTTGAACGTGCAGAGTGAATGGTGGTGAATGATG 1489  
QY 61 CTGATCGTTACTGTGTAACGCGCCACAGTGTATCTCCGGAAGCGCGTCCCGTG 120  
DB 1488 CTGATCGTTACTGTGTAACGCGCCACAGTGTATCTCCGGAAGCGCGTCCCGTG 1429  
QY 121 GTTAAAGTAAATCCATCGAAGAACTCCGGGGCGCGCTTAACTGGCGATGATATC 180  
DB 1428 GTTAAAGTAAATCCATCGAAGAACTCCGGGGCGCGCTTAACTGGCGATGATATC 1369  
QY 181 GCTTCTCGGAGTATGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 1368 GCTTCTCGGAGTATGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309  
QY 241 GCGCTAGTAATCTGTGCGCAAGTCAAGTCAATGCGAATCTGCTTCTGTACGACG 300  
DB 1308 GCGCTAGTAATCTGTGCGCAAGTCAAGTCAATGCGAATCTGCTTCTGTACGACG 1249  
QY 301 CATCCGACCATTCGAAATTAACGGGTACTTCCCGCAACCAAGCTGATCCCTCGAT 360  
DB 1248 CATCCGACCATTCGAAATTAACGGGTACTTCCCGCAACCAAGCTGATCCCTCGAT 1189  
QY 361 TTGGAAGAGGTTTCAAGAGGTGTGATCCGACGCTGCAACGAGGATTAATCAGCG 420  
DB 1188 TTGGAAGAGGTTTCAAGAGGTGTGATCCGACGCTGCAACGAGGATTAATCAGCG 1129  
QY 421 CTGAGTTCAATTTGCGCGCTGTGCTTCTGACTAGCCCAAGTGCCTGGCAAGCTA 480  
DB 1128 CTGAGTTCAATTTGCGCGCTGTGCTTCTGACTAGCCCAAGTGCCTGGCAAGCTA 1069  
QY 481 CAGCAGATGATCCACTGCGCGCTTAAAGCGGGTGTCCGCTGCTGATTCAAAAGT 540  
DB 1068 CAGCAGATGATCCACTGCGCGCTTAAAGCGGGTGTCCGCTGCTGATTCAAAAGT 1009  
QY 541 ACCGATTTGAGGCTTACCGCGCGCGCTGCTGCTTAAAGCGGATCTCTCGAATTTGAA 600  
DB 1008 ACCGATTTGAGGCTTACCGCGCGCGCTGCTGCTTAAAGCGGATCTCTCGAATTTGAA 949  
QY 601 GCTGTGTCGCTAAATGTAAGACGGAAGAGATTTGAGCGCGCATGAAATCGATT 660  
DB 948 GCTGTGTCGCTAAATGTAAGACGGAAGAGATTTGAGCGCGCATGAAATCGATT 989  
QY 661 GCCGATTAAGAACTCTCGGCTCTGTTAGTAAACCCGTTCCGAACAAGGATATGCTGCTG 720  
DB 888 GCCGATTAAGAACTCTCGGCTCTGTTAGTAAACCCGTTCCGAACAAGGATATGCTGCTG 829  
QY 721 CAACCGGTTAAAGCGCGCTGCTATATGCCAACCAAGCGCAGAGATGATGCTTACC 780  
DB 828 CAACCGGTTAAAGCGCGCTGCTATATGCCAACCAAGCGCAGAGATGATGCTTACC 769  
QY 781 GGTGCGGCGCAACGCTGATTTGCGCTCTGTGCGGCAACGCTGCAAGCGGTAATTCGCTG 840  
DB 768 GGTGCGGCGCAACGCTGATTTGCGCTCTGTGCGGCAACGCTGCAAGCGGTAATTCGCTG 709

QY 841 GAAGAGCTCTCTTTGCAATGCGGCGCTGCGGCTGCTGCGCAAACTGGAAACC 900  
DB 708 GAAGAGCTCTCTTTGCAATGCGGCGCTGCGGCTGCTGCGCAAACTGGAAACC 649  
QY 901 TCACGCGTTTCCGCGATCGAGCTGAAATGCTGTACGTGACGTGACATGAGCTTT 960  
DB 648 TCACGCGTTTCCGCGATCGAGCTGAAATGCTGTACGTGACGTGACATGAGCTTT 589  
QY 961 GCGCGTATGACCGAAGGAACTGAGCTGCGGCTGAGCGGAGCGCTAAACGTGAGAA 1020  
DB 588 GCGCGTATGACCGAAGGAACTGAGCTGCGGCTGAGCGGAGCGCTAAACGTGAGAA 529  
QY 1021 AAGTGTGATGATCAACAGCTGTCTTTGACATCTCTGACGCGCGGACGCTCTTATCTG 1080  
DB 528 AAGTGTGATGATCAACAGCTGTCTTTGACATCTCTGACGCGCGGACGCTCTTATCTG 469  
QY 1081 GCAATGCGCGCAAGCTGAGTGAACGCTGTGATGTTGCGGTCAACAGCGATGCTCCACC 1140  
DB 468 GCAATGCGCGCAAGCTGAGTGAACGCTGTGATGTTGCGGTCAACAGCGATGCTCCACC 409  
QY 1141 AAGCGCTGAAAGGAGATTCGCGCGCGGTAAACCACTCGAACAAGCTATGATGCTG 1200  
DB 408 AAGCGCTGAAAGGAGATTCGCGCGCGGTAAACCACTCGAACAAGCTATGATGCTG 349  
QY 1201 GCGCGACTGAAAGCGCTGCACTGAGTGTCTTTGAAAGAGACAGCGCGCAGCGCTTG 1260  
DB 348 GCGCGACTGAAAGCGCTGCACTGAGTGTCTTTGAAAGAGACAGCGCGCAGCGCTTG 289  
QY 1261 ATGCGCGGAGATCTTGGCAAGATCTGCTGTGAAAGCGCGCGCACTATTAACCAAGAGATT 1320  
DB 288 ATGCGCGGAGATCTTGGCAAGATCTGCTGTGAAAGCGCGCGCACTATTAACCAAGAGATT 229  
QY 1321 GCGCGAGTAAAGAGTCTGCGGCGCAACGCTGCGCAAGCTGTGAGCTCAACTTTGAAGAC 1380  
DB 228 GCGCGAGTAAAGAGTCTGCGGCGCAACGCTGCGCAAGCTGTGAGCTCAACTTTGAAGAC 169  
QY 1381 GGTGTCTGACGACCAATCATCATCAAGAGATCCACAGATTAAGAGCTAA 1434  
DB 168 GGTGTCTGACGACCAATCATCATCAAGAGATCCACAGATTAAGAGAGCTAA 115

RESULT 7  
U00096.31/c  
WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1100000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000

U00096\_26 2600001 2710000  
U00096\_27 2700001 2810000  
U00096\_28 2800001 2910000  
U00096\_29 2900001 3010000  
U00096\_30 3000001 3110000  
U00096\_31 3100001 3210000  
U00096\_32 3200001 3310000  
U00096\_33 3300001 3410000  
U00096\_34 3400001 3510000  
U00096\_35 3500001 3610000  
U00096\_36 3600001 3710000  
U00096\_37 3700001 3810000  
U00096\_38 3800001 3910000  
U00096\_39 3900001 4010000  
U00096\_40 4000001 4110000  
U00096\_41 4100001 4210000  
U00096\_42 4200001 4310000  
U00096\_43 4300001 4410000  
U00096\_44 4400001 4510000  
U00096\_45 4500001 4610000  
U00096\_46 4600001 4639675

Continuation (32 of 47) of U00096 from base 3100001 (U00096 Escherichia coli K-12 MG1655)

Query Match 100.0%; Score 1434; DB 1; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGTAAACGCTGCGCAGGTTTGAACGTGACGAGATGATGCTGATGATGATG 60  
DB 94775 ATGAAGTAAACGCTGCGCAGGTTTGAACGTGACGAGATGATGATGATGATG 94716  
QY 61 CTGATCGTTTACGTGATGACGCGCCACAGTCTATCTGCGCGAAGCGCGGTGCCGTG 120  
DB 94715 CTGATCGTTTACGTGATGACGCGCCACAGTCTATCTGCGCGAAGCGCGGTGCCGTG 94656  
QY 121 GTTAAAGTAAATCCATCGAAGAGCTCGGCGCGCGCTTACGTGCGATGATGATTC 180  
DB 94655 GTTAAAGTAAATCCATCGAAGAGCTCGGCGCGCGCTTACGTGCGATGATGATTC 94596  
QY 181 GCTTCTCTGCGGCTATGACGCGCTGCTGCGGTTGACGCGGCAATGACATGACGCGC 240  
DB 94595 GCTTCTCTGCGGCTATGACGCGCTGCTGCGGTTGACGCGGCAATGACATGACGCGC 94536  
QY 241 GCGCTGAGTAAATCTGCGCGCAGCTCAACGTCAAAATGCGACTTCTGTACCGACG 300  
DB 94535 GCGCTGAGTAAATCTGCGCGCAGCTCAACGTCAAAATGCGACTTCTGTACCGACG 94476  
QY 301 CATCCGACATTAACCAATTAACGCGGTAATCTTCCGCGCAACCAAGCTGATCCGTGAT 360  
DB 94475 CATCCGACATTAACCAATTAACGCGGTAATCTTCCGCGCAACCAAGCTGATCCGTGAT 94416  
QY 361 TTTGAAGAGGTTTCAAGGTTGATCCGACGCGCTGACAGCGGATGATTAATCAGGCG 420  
DB 94415 TTTGAAGAGGTTTCAAGGTTGATCCGACGCGCTGACAGCGGATGATTAATCAGGCG 94356  
QY 421 CTGAGTTGATGCGCGCTGCTGCTTCTGACTGACGCAACCAAGGTGCGTGCAGAGCTA 480  
DB 94355 CTGAGTTGATGCGCGCTGCTGCTTCTGACTGACGCAACCAAGGTGCGTGCAGAGCTA 94296  
QY 481 CAGCAGATGATCCACGCGCGCTGAAAGCGGCTGCTCCGCTGATGATCCAAAAGGT 540  
DB 94295 CAGCAGATGATCCACGCGCGCTGAAAGCGGCTGCTCCGCTGATGATCCAAAAGGT 94236  
QY 541 ACCGATTTGAGGCTACCGCGCGCGCTACGCTGTAACGCGCAATCTCGGAATTTGAA 600  
DB 94235 ACCGATTTGAGGCTACCGCGCGCGCTACGCTGTAACGCGCAATCTCGGAATTTGAA 94176  
QY 601 GCTGTTGCTGATTAATGTAAGACCGAAGAGATGTTGAGCGCGCATGAATGAT 660  
DB 94175 GCTGTTGCTGATTAATGTAAGACCGAAGAGATGTTGAGCGCGCATGAATGAT 94116  
QY 661 GCCGATTAAGACTCTGCGCTCTGTTAGTACCCGTTCCGAACAGGATATGCTGCTG 720



QY 421 CTGAGTTGATTTGGCGGCTGGTCTTCTGACTACGCCAAGGTGGCTGGCAAGCTA 480  
DB 103176 CTGAGTTGATTTGGCGGCTGGTCTTCTGACTACGCCAAGGTGGCTGGCAAGCTA 103117  
QY 481 CAGCAGATGATTCGACCTGGCGGCTGTAAGCGGGTGTTCGGGTCTGATTCGAAAGGT 540  
DB 103116 CAGCAGATGATTCGACCTGGCGGCTGTAAGCGGGTGTTCGGGTCTGATTCGAAAGGT 103057  
QY 541 ACCGATTTTGGACGCTACCGCGGCGCTACGCTGTTAAAGCGGATCTCTCGGAATTTGAA 600  
DB 103056 ACCGATTTTGGACGCTACCGCGGCGCTACGCTGTTAAAGCGGATCTCTCGGAATTTGAA 102997  
QY 601 GCTGTTGTCTGTAATATGTAAGACCGAAGAGATTTGTAAGCGGCGCATGAACTGATT 660  
DB 102996 GCTGTTGTCTGTAATATGTAAGACCGAAGAGATTTGTAAGCGGCGCATGAACTGATT 102937  
QY 661 GCCGATTAACGAACTCTCGGCTCTGTAGTACCCGTTCCGAAACGGGTATTCCTGCTG 720  
DB 102936 GCCGATTAACGAACTCTCGGCTCTGTAGTACCCGTTCCGAAACGGGTATTCCTGCTG 102877  
QY 721 CAACCGGGTAAAGGCGGCTGATATGCAACCAAGCGGAGAAAGTATGACGTTACC 780  
DB 102876 CAACCGGGTAAAGGCGGCTGATATGCAACCAAGCGGAGAAAGTATGACGTTACC 102817  
QY 781 GGTGCGGGCGACACGCTGATTTGGCGTCTGCGGCAACGCTGCGACGGGGTATTCGCTG 840  
DB 102816 GGTGCGGGCGACACGCTGATTTGGCGTCTGCGGCAACGCTGCGACGGGGTATTCGCTG 102757  
QY 841 GAAGAAGCTGCTCTTTTCCAAATGCGCGGCTGCGGTGTGTGCGGCAACCTGGGAACC 900  
DB 102756 GAAGAAGCTGCTCTTTTCCAAATGCGCGGCTGCGGTGTGTGCGGCAACCTGGGAACC 102697  
QY 901 TCACGCGTTTGGCGGATCGACTGAAATATCTGTACGTGACGTGACATACAGGCTTT 960  
DB 102696 TCACGCGTTTGGCGGATCGACTGAAATATCTGTACGTGACGTGACATACAGGCTTT 102637  
QY 961 GCGCTGATGACCGAAGAGAACTGAAGCTGCGCTGACGCGGCAAGCTGTTAA 1020  
DB 102636 GCGCTGATGACCGAAGAGAACTGAAGCTGCGCTGACGCGGCAAGCTGTTAA 102577  
QY 1021 AAAGTGTGATGACCAACGCTGTTTGAACATCTCTGACGCGGCGCAAGCTCTTATCTG 1080  
DB 102576 AAAGTGTGATGACCAACGCTGTTTGAACATCTCTGACGCGGCGCAAGCTCTTATCTG 102517  
QY 1081 GCAATGCGCGCAAGCTGCTGATGACGCTGTTGTCGCTGCAACAGGATGCTCCACC 1140  
DB 102516 GCAATGCGCGCAAGCTGCTGATGACGCTGTTGTCGCTGCAACAGGATGCTCCACC 102457  
QY 1141 AAACGGCTGAAAGGGGATTCGCGCCGCTAAACCACTCGAACAGCGTATGATTGCTG 1200  
DB 102456 AAACGGCTGAAAGGGGATTCGCGCCGCTAAACCACTCGAACAGCGTATGATTGCTG 102397  
QY 1201 GGGCGACTGGAAGGGGCTGACTGGGTAGTGTGTTGAAGGAGCAAGCGGCAAGCGCTG 1260  
DB 102396 GGGCGACTGGAAGGGGCTGACTGGGTAGTGTGTTGAAGGAGCAAGCGGCAAGCGCTG 102337  
QY 1261 ATCCCGCGGATCTTGCCAGATCTGCTGTGTAAGAGCGGCGACTATAACGAGAAGATT 1320  
DB 102336 ATCCCGCGGATCTTGCCAGATCTGCTGTGTAAGAGCGGCGACTATAACGAGAAGATT 102277  
QY 1321 GCCCGAGTAAAGAAAGTCTGGGCGCAACGCTGGCAAGTGTGTGCTCAACTTTGAAAGC 1380  
DB 102276 GCCCGAGTAAAGAAAGTCTGGGCGCAACGCTGGCAAGTGTGTGCTCAACTTTGAAAGC 102217  
QY 1381 GGTGTCTGAGAGCAACATCATCAAGAGATTCGACGAGTAAAGAGCTAA 1434  
DB 102216 GGTGTCTGAGAGCAACATCATCAAGAGATTCGACGAGTAAAGAGCTAA 102163

RESULT 9  
AB005174.40/c  
WPCOMMENT  
Sequence split into 56 fragments LOCUS AB005174 Accession AB005174

Fragment Name Begin End  
AB005174.00 1 110000  
AB005174.01 100001 210000  
AB005174.02 200001 310000  
AB005174.03 300001 410000  
AB005174.04 400001 510000  
AB005174.05 500001 610000  
AB005174.06 600001 710000  
AB005174.07 700001 810000  
AB005174.08 800001 910000  
AB005174.09 900001 1010000  
AB005174.10 1000001 1110000  
AB005174.11 1100001 1210000  
AB005174.12 1200001 1310000  
AB005174.13 1300001 1410000  
AB005174.14 1400001 1510000  
AB005174.15 1500001 1610000  
AB005174.16 1600001 1710000  
AB005174.17 1700001 1810000  
AB005174.18 1800001 1910000  
AB005174.19 1900001 2010000  
AB005174.20 2000001 2110000  
AB005174.21 2100001 2210000  
AB005174.22 2200001 2310000  
AB005174.23 2300001 2410000  
AB005174.24 2400001 2510000  
AB005174.25 2500001 2610000  
AB005174.26 2600001 2710000  
AB005174.27 2700001 2810000  
AB005174.28 2800001 2910000  
AB005174.29 2900001 3010000  
AB005174.30 3000001 3110000  
AB005174.31 3100001 3210000  
AB005174.32 3200001 3310000  
AB005174.33 3300001 3410000  
AB005174.34 3400001 3510000  
AB005174.35 3500001 3610000  
AB005174.36 3600001 3710000  
AB005174.37 3700001 3810000  
AB005174.38 3800001 3910000  
AB005174.39 3900001 4010000  
AB005174.40 4000001 4110000  
AB005174.41 4100001 4210000  
AB005174.42 4200001 4310000  
AB005174.43 4300001 4410000  
AB005174.44 4400001 4510000  
AB005174.45 4500001 4610000  
AB005174.46 4600001 4710000  
AB005174.47 4700001 4810000  
AB005174.48 4800001 4910000  
AB005174.49 4900001 5010000  
AB005174.50 5000001 5110000  
AB005174.51 5100001 5210000  
AB005174.52 5200001 5310000  
AB005174.53 5300001 5410000  
AB005174.54 5400001 5510000  
AB005174.55 5500001 5528445

Continuation (41 of 56) of AB005174 from base 4000001 (AB005174 Escherichia coli O157:H7)

Query Match 96.8%; Score 1387.6; DB 1; Length 110000;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 1405; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 ATGAAGTAAAGCTGCGCAAGTTTGAACGTGACGAGATGATGGTGGTGAATGATG 60  
DB 3596 ATGAAGTAAAGCTGCGCAAGTTTGAACGTGACGAGATGATGGTGGTGAATGATG 3537  
QY 61 CTGGAATGTTACTGTGTAAGCGCCACCAAGTGTATCTCGCCGGAAGCGCGGCTCCGCTG 120  
DB 3536 CTGGAATGTTACTGTGTAAGCGCCCTCAAGTGTATCTCGCCGGAAGCGCGGCTCCGCTG 3477  
QY 121 GTTAAAGTAAATCAATCGAAGAACTCCGGGCGCGCGCTTAACTGGGATGAATATC 180





BA000007 54 5400001 5498450  
Continuation (40 of 55) of BA000007 from base 3900001 (BA000007 Escherichia coli O157:H7)



```

/codon_start=1
/translate_table=11
/product="Hypothetical protein yqkC"
/protein_id="PAN82241.1"
/db_xref="GI:26110055"
/translation="MTYPRRTMALNTYYQRETTMDPKKIBQIAROVHSMKPIRE
FGDVEKKIRQTLQAOLTRDLVSRBSFDVQTVLITRKALLLEQRISELBARVNP
VALQSPRAIPPDYKAE"
complement(582..788)
/gene="g1g5"
/locus_tag="c3797"
/complement(582..788)
/gene="g1g5"
/locus_tag="c3797"
/functions="putative enzyme; Macromolecule synthesis,
modification: Polysaccharides - (cytoplasmic)"
/note="Escherichia coli K-12 ortholog: b3049; Escherichia
coli O157:H7 ortholog: z4401"
/codon_start=1
/translate_table=11
/product="Glycogen synthesis protein g1gS"
/protein_id="PAN82242.1"
/db_xref="GI:26110056"
/translation="MLMDHSLNSLNFDFLARSFARMHAGRPVDILAVTGNDEHR
TWFCARYAWYCCQWMTRELEHL"
1052..1681
/gene="yqkU"
/locus_tag="c3798"
1052..1681
/gene="yqkU"
/locus_tag="c3798"
/functions="putative"
/note="Escherichia coli K-12 ortholog: b3050; Escherichia
coli O157:H7 ortholog: z4402"
/codon_start=1
/translate_table=11
/product="Hypothetical protein yqkU"
/protein_id="PAN82243.1"
/db_xref="GI:26110057"
/translation="WLFADYNTPEYLFALFISFVLLIGLLETIPALICGHTLSGALDHLID
HYDSITTGHSIQALMYLNIIGRLPALVVLCLAGFPGLIGILQHCATLWQSPINLF
VDSVSLFTIIVAVYTGKVVAPMIPRHSASITEEBYGSMALITGHQATGNCPEKG
LTQDFQIHYLLLEPEBEGKFTKGDVLIICRLSATRYLLNNPWMSGL"
1708..3369
/gene="yqkV"
/locus_tag="c3799"
1708..3369
/gene="yqkV"
/locus_tag="c3799"
/functions="putative membrane"
/note="Escherichia coli K-12 ortholog: b3051; Escherichia
coli O157:H7 ortholog: z4403"
/codon_start=1
/translate_table=11
/product="Hypothetical protein yqkV"
/protein_id="PAN82244.1"
/db_xref="GI:26110058"
/translation="WDDIVNSVPSMFTAILAVCILFIIITPARLYRRASQAQVFR
TGLSGQVNSGAIWPIFIHBIIPIMNTLKLKVSRTIDSLITTDKRMVDVVAF
VRVPSVEGIATAQTLGORTLSPEIDRLMVEDFVNLPAQMTMHELOTRERF
VOGQONTVAADLSKNGLEBSVSLTNFNQSKHFNPNNAFDABGTLKLOETERRR
ERRNEVDVAVAREKRDALSKLETLEOQAPPTLEOQOVTTRTAEONAKTAAAPA
ERRRQAQTIILERQIQETETIDREQAVNSQVNETTROQAEADRAQVALIAAQDAETK
ITALAKSEQSQAEARALALAEVNSAQVNETTROQAEADRAQVALIAAQDAETK
FALITVALAKAEKAEWQAAIYELIATRKGLABEAQRALANDAINVLSIDQTSIK
FKALIQALPAVIEKSVPMKAIIGIKITIQDGLNRGATGADANTGVGGNLAEQLL
SALSVRTQALPILDSLNEIGVSGGSLAAL7TSLSSTPTVEKAE"
complement(3788..5221)
/locus_tag="c3800"
complement(3788..5221)
/locus_tag="c3800"
/BC_number="2.7.-.-"
/function="putative"

```

```

/note="Escherichia coli K-12 ortholog: b3052; Escherichia
coli O157:H7 ortholog: z4405"
/codon_start=1
/translate_table=11
/product="ADP-heptose synthase"
/protein_id="PAN82245.1"
/db_xref="GI:26110059"
/translation="MKVTLPEPERAGVWVVDVMDLRVYWGPTSRISPEPVVYKN
TTERPGAAANYAMNIASIGANRLVGLTGIDDAARALSKSLADVNVKDFVSVPH
TITKLVLNRNQQLLRLDPFEGEGVDPPQRLPFRNIVRAQVLAALVTSYKALASV
QQMIOIARAGVPLIDPKGTDFERYRAGTLLTPNISREPAVVGKCTEELVERKMK
LIADYLSALVTRSEQWSLQPGKAPLHPPOAEVDVVGAGDTVGVLAATLAA
GNSLEBACFPANAAAGVVGKGTSTVSPILENAVRGADDFGVMTSEELAAVA
ARRKGRVMTNGVDPILHAGHSYLANRKGDRLIVVNSPASTKRLKGDSPVNP
LEQRMVILGALREVDVVSPEBDTPQRLAGILPDLVYNGDVKPEBISAKSEVANG
GEVLVNFEGGCTNIIYIKIQDKRG"
complement(569..8220)
/gene="g1n5"
/locus_tag="c3801"
complement(569..8220)
/gene="g1n5"
/locus_tag="c3801"
/BC_number="2.7.7.42"
/functions="enzyme; Proteins - translation and
modification"
/note="Escherichia coli K-12 ortholog: b3053; Escherichia
coli O157:H7 ortholog: z4406"
/codon_start=1
/translate_table=11
/product="Glutamate-ammonia-lyase adenylyltransferase"
/protein_id="PAN82246.1"
/db_xref="GI:26110060"
/translation="MRGNALKNISVWGQTIIRNSGCTAENVNOARIKIMKPLSP
LQYWTVAVERLEPPLAESLSAQASVITFSDPFVDSVTAHPEMTLEBSQPDQE
WQHYASWMLQBALSNVSDGALMRELRFRRIINVRIAMQTLALVTESLLOQLSHA
ETLIVARDPMLYDACCEBWTGCPNAGQEPVLIIDMGKLGGLNLPSSDIDIFAMP
BHCCTGGRRELDNAPFTFMGQRLIKVDPQDPQDVFVRVMDRLPFGSGPLVFSF
ALIEDYVOGGRMERVYAWYKARIMGDSGVYANVELRAMLPRPFPRYIDFVYQSLR
NMKGMIARBYRRRGITDNPVKTGACGIRIETFYVQVQLRGCEBPLQSRALLPTLSA
IAPHLHLSNDABQLAVATLFLRRLNLSQINDBOTQTLPPBELRARALAMMDPAD
WPDQTKOVLTAHMANVRVNEELIGDSESTQBSLSQRELMQDLOEDTVPVLAH
LESDPDLTVAHADRKEIDKRTTGTPRSGOVLDHIMPHLSQVCARBAVAVLSIT
ALLGVITRTYVLELSERPALKHLISCAASPMASOLARPLLDLDELDPNTLYQ
PATDVAVERLEIROYLIRVPEDEDEQQLALROPKQOLRIAAADTAGTLPVKVSQD
LITLARMIDAVYQAAVQWVARYGCPNHLNREBGRFAYVNGYCGKCGEILGYSSDL
LIFHDCEPMDATDGERIDGROFYRLRQRIWHLSTRTSSGILYVDARLAPSGAA
GMLVTSAEAFADYQKNBAWTHWQALVRAVVGDPQLTAHPDAVAREITLTPREGKT
LOTEVEREMERKMAHLGNKHDRFDIKADGGITDIEFTQVYLVAREAHKPKLTERMS
DNVRIELEAONDIMEBQAMALTRAVYTLRLDELHLALQELRGVHSBDCFALRELV
RASQKMLVER"
complement(8132..9460)
/gene="yqkF"
/locus_tag="c3802"
complement(8132..9460)
/gene="yqkF"
/locus_tag="c3802"
/note="TOR2X2; Escherichia coli K-12 ortholog: b3054;
Escherichia coli O157:H7 ortholog: z4407"
/codon_start=1
/translate_table=11
/product="Hypothetical protein yqkF"
/protein_id="PAN82247.1"
/db_xref="GI:26110061"
/translation="WGVPSGGMAQIEIKPTVNSAVBALRDHLNTLOGEHRDVPQ
LNIIVETPDNMDRGDMGLRIRGENGRLEMTKVAVGTGGLHQEPENVALSEPL
DLAQLPTEWPNELPADLASRVQPLSFDPYAEKMLVVDGSRIBIADPOGVKGE
FAPBICELEIELLSGDMRAVLKLANQLVSGTGRQGLSVDAAGVLAQGNPAREIKP
TTIIVAAKADVQGLKALAEALANOMYHEBLMWGNSAAKQVLAAGLVHTLMU
FGGIVPRKASTHRLDLLTOCEATIASVFCQPLGDRDQLPRLITDIBTLLILA
LKKKADSKTSDSFPRADTHLSRHAAELKSVFCQPLGDRDQLPRLITDIBTLLILA
GYDDPVAAQDWLWNGQLRHAIAITGQRIETIEHRNANNQEPFWLHSGKR"
complement(9503..9955)
/locus_tag="c3803"

```





CDS  
 /locus\_tag="B3037"  
 complement(7274..9520)  
 /gene="ptp"  
 /locus\_tag="B3037"  
 /function="enzyme; Transport of small molecules: Other"  
 /note="Residues 1 to 736 of 748 are 99.59 pct identical to  
 residues 1 to 736 of 748 from Escherichia coli K-12 ;  
 B2829"  
 /codon\_start=1  
 /transl\_table=1  
 /product="PTS system, enzyme I, transcriptional regulator  
 (with NTP and NTP proteinase)"  
 /protein\_id="AAP1815.1"  
 /db\_xref="gi:30042427"

/translation="MLTRIREIVKVASAPRLNALNLTVDICLADTEVCVYLLD  
 HDRCYLLMTRGKAKRRGRTVTLAPBEGVGLVGRAPRLNADAKHSPKTPSY  
 KEERFRAVLGVPIIQRRQLGVLVQREMLQDSESSPLVLTATMALISQSLT  
 ALFQYRQTRIRALPAPVAVIAGMODATLPLMEQVYQASLTDPLERBLTGLE  
 AANFRRYSKRFAPAGQKETAIFDLVSHLSIDRLRELFADVDKGSVMAVATYI  
 EKAEPAPALSNDYLKERAGDRLALGQRLPHLDADQNPAPERFLVADLSATT  
 LAELPDRIQVGVVVDGANSNAIIMRALGIPVMGADIDPSVLRHTLIDVGRGE  
 LILDEPVLQHYRQLISERIEISRLABDDVNLPAQKSGERITVMLNAGLSPEHEK  
 LQSCIDELGTRIPMLQSPSPSESVQAVQIQMLQMDKPVTLTLDVGLDKL  
 PYMPISENPCLGWRGIRITLDQEIFLIQVRAMLRANAAATGNILPLVTSIDED  
 EARLIRAGREVEEMIGYEIPKPRIGIMLEPVMPLPLARVDFISVGTDLQ  
 YILAVDRNTRVANIYDSLHPAMRLALAMTAREARIGHIDLRCGEVADPWCVALDI

Query Match 95.9%; Score 1374.8; DB 1; Length 292906;  
 Best Local Similarity 97.4%; Pred. No. 6.9e-311;  
 Matches 1397; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

1 ATGAAAGTAACTGCGCAGAGTTTGAACGTGACAGAGTATGATGATGATGATG 60  
 260584 ATGAAAGTAACTGCGCAGAGTTTGAACGTGACAGAGTATGATGATGATG 260525  
 61 CTGATCGTTACTGATGAGCGCCCAACAGTGTATCTTCCCGAAGCGCGTCCG 120  
 260524 CTGATCGTTACTGATGAGCGCCCAACAGTGTATCTTCCCGAAGCGCGTCCG 260465  
 121 GTTAAAGTAACTGCGCAGAGTTTGAACGTGACAGAGTATGATGATGATGATG 180  
 260464 GTTAAAGTAACTGCGCAGAGTTTGAACGTGACAGAGTATGATGATGATGATG 260405  
 181 GCTTCTCTCGGTGATGACAGCGCTGGTGGTGAAGCGGCAATTGACAGTACAG 240  
 260404 GCTTCTCTCGGTGATGACAGCGCTGGTGGTGAAGCGGCAATTGACAGTACAG 260345  
 241 GCGCTGAGTAAATCTCTGCGCGACGTCAACGTCAATGCGACTTCTGTTCTGA 300  
 260344 GCGCTGAGTAAATCTCTGCGCGACGTCAACGTCAATGCGACTTCTGTTCTGA 260285  
 301 CATCGACCAATTACCAATTACGGGTACTTCCCGCAACCAACAGCTGATCCG 360  
 260284 CATCGACCAATTACCAATTACGGGTACTTCCCGCAACCAACAGCTGATCCG 260225  
 361 TTGTAAGAGGTTTGAAGTGTGATCCGACGCGCTGACAGCGGATTAATCAGCG 420  
 260224 TTGTAAGAGGTTTGAAGTGTGATCCGACGCGCTGACAGCGGATTAATCAGCG 260165  
 421 CTGAGTTGATGCGCGCTGCTGCTTCTGATCTACGCGCAAAAGTGCGCTGCA 480  
 260164 CTGAGTTGATGCGCGCTGCTGCTTCTGATCTACGCGCAAAAGTGCGCTGCA 260105  
 481 CAGAGAGTATCAACGCGCGCTGAAAGCGGTGTTCCGGGTGATGATGATGATG 540  
 260104 CAGAGAGTATCAACGCGCGCTGAAAGCGGTGTTCCGGGTGATGATGATGATG 260045  
 541 ACCGATTTGAGCGCTACCGCGCGCTACGCTGTTTAAACGCGGAATCTCTCGA 600  
 260044 ACCGATTTGAGCGCTACCGCGCGCTACGCTGTTTAAACGCGGAATCTCTCGA 259985  
 601 GCTGTTGCGTAAATGTAAGCCGAGAGAGATGTTGAGCGCGGCTGAAATCTGAT 660

259984 GCTGTTGCGTAAATGTAAGCCGAGAGAGATGTTGAGCGCGCATGAACTGAT 259925  
 661 GCCGATTTAGCACTCTCGGCTCTGTTAGTACCCGTTCCGACAGGGTATGTCGTC 720  
 259924 GCCGATTTAGCACTCTCGGCTCTGTTAGTACCCGTTCCGACAGGGTATGTCGTC 259865  
 721 CAACCGGGTAAAGCGCGCTGATATGCCAACCCGACGAGAGTATGATGATGAT 780  
 259864 CAACCGGGTAAAGCGCGCTGATATGCCAACCCGACGAGAGTATGATGATGAT 259805  
 781 GGTGCGCGGACACACGCTGATGCTCTGTCGCGGACACGCTGACGCGGTAATG 840  
 259804 GGTGCGCGGACACACGCTGATGCTCTGTCGCGGACACGCTGACGCGGTAATG 259745  
 841 GAAGAAGCCGCTCTCTTGGCAATGCGCGCGCTGCGGTGATGCGCAACTGGAA 900  
 259744 GAAGAAGCCGCTCTCTTGGCAATGCGCGCGCTGCGGTGATGCGCAACTGGAA 259685  
 901 TCACGCTTTCGCGGATCGGATCGAAGTGAATAATGCTGACGTGACATACAGCT 960  
 259684 TCACGCTTTCGCGGATCGGATCGAAGTGAATAATGCTGACGTGACATACAGCT 259625  
 961 GCGGTGATGACCGGAGAGAACTGAAGCTGCGCGCTGACCGGCAAGCTGATG 1020  
 259624 GCGGTGATGACCGGAGAGAACTGAAGCTGCGCGCTGACCGGCAAGCTGATG 259565  
 1021 AAGGTGATGACCGGAGAGAACTGAAGCTGATGATGATGATGATGATGATGAT 1080  
 259564 AAGGTGATGACCGGAGAGAACTGAAGCTGATGATGATGATGATGATGATGAT 259505  
 1081 GCAATGCGCGGACGAGTGGTGAACGCTGATGATGATGATGATGATGATGATG 1140  
 259504 GCAATGCGCGGACGAGTGGTGAACGCTGATGATGATGATGATGATGATGATG 259445  
 1141 AAACGCTGAAGAGGATTCGCGCGCGGATTAACCACTGCAACAGCTATGATG 1200  
 259444 AAACGCTGAAGAGGATTCGCGCGCGGATTAACCACTGCAACAGCTATGATG 259385  
 1201 GCGGCACTGGAACGCTGCACTGAGTATGATGATGATGATGATGATGATGATG 1260  
 259384 GCGGCACTGGAACGCTGCACTGAGTATGATGATGATGATGATGATGATGATG 259325  
 1261 ATGCGCGGATCTTGCAGATCTGCTGTAAGAGGCGGCGCATTAACCAAGAG 1320  
 259324 ATGCGCGGATCTTGCAGATCTGCTGTAAGAGGCGGCGCATTAACCAAGAG 259265  
 1321 GCGGAGATTAAGAGTCTGCGGCAACGCTGCGGAGTGTGCTGCTCACTTTGA 1380  
 259264 GCGGAGATTAAGAGTCTGCGGCAACGCTGCGGAGTGTGCTGCTCACTTTGA 259205  
 1381 GGTGCTGACGACCAACATCATCAAGAGATCCACAGGATTAAGAGCTTA 1434  
 259204 GGTGCTGACGACCAACATCATCAAGAGATCCACAGGATTAAGAGCTTA 259151

RESULT 13  
 AB005674\_31/c  
 WPCOMMENT  
 Sequence split into 46 fragments LOCUS AB005674 Accession AB005674

Fragment Name	Begin	End
AB005674_00	1	11000
AB005674_01	10001	21000
AB005674_02	20001	31000
AB005674_03	30001	41000
AB005674_04	40001	51000
AB005674_05	50001	61000
AB005674_06	60001	71000
AB005674_07	70001	81000
AB005674_08	80001	91000
AB005674_09	90001	101000
AB005674_10	100001	111000
AB005674_11	110001	121000
AB005674_12	120001	131000







Details of *S. typhi* sequencing at the Sanger Centre are available on the World Wide Web. ([http://www.sanger.ac.uk/Projects/S\\_typhi/](http://www.sanger.ac.uk/Projects/S_typhi/)).

## FEATURES

## source

```

1. 258050
/organism="Salmonella enterica subsp. enterica serovar Typhi"
/mol_type="genomic DNA"
/strain="CT18"
/db_xref="taxon:90370"
69. 1544
/gene="tolC"
/note="synonym: STY3364"
69. 1544
/gene="tolC"
/note="Similar to Salmonella enteritidis outer membrane protein TolC precursor tolC SW:tolC_SALBN (Q54001) (491 aa) fasta scores: E(): 0, 99.04 id in 491 aa, and to Escherichia coli outer membrane protein tolC precursor tolC or mcb or mka or refi SW:tolC_ECOLI (P02930) (495 aa) fasta scores: E(): 0, 89.74 id in 495 aa Orthologue of E. coli tolC (TOLC_ECOLI); Fasta hit to tolC_ECOLI (495 aa), 90% identity in 495 aa overlap"
/codon_start=1
/transl_table=11
/product="outer membrane protein TolC precursor"
/protein_id="CAD07712.1"
/db_xref="GI:16504264"
/db_xref="GOA:Q823N8"
/db_xref="InterPro:IPR003423"
/db_xref="InterPro:IPR010130"
/db_xref="UniProt/TREMBL:Q823N8"
/transl_table=11
/translation="MOKKLLPITIGSLGFSYTLGAENTLMQYVQARLSNPELRS AARDDAPEKINERASPLLPOLGLGADYTSNGRDNGINSNETSASLDLTGLFPM SKRGLTLQERKAGIDVTYQTDQDTLLNTANAYFVLAIDVLSTQAKAIYQO LDOTGRFNVGLVAITDVQNAQAYDTLVANETVARNLDNAVELEIQVQNYBELA SLAVEHFKTKPKVNALEKEAENRNLILQARLSODLAREQIRQADGHLPTNLTRA STGISDTSVSGSKTNSAQVDSNMGKIGLNPGLPLVQSGMNSQVKAQVAVGAS EQLBSARISVYQTRSSFPNNINASSINNAIKQAVVQAQSSLDAMEAGYVGRITTYD VLDAITTLVDKQQLANARYTYLLNQINITYALGTLNEQDLALNLTGLKPIPTSPS VAPETPEODAAADQYNHSAAPAVQPTAANANSNGNPFRR"
1757. 2428
/gene="STY3365"
1757. 2428
/gene="STY3365"
/note="Orthologue of E. coli ygiB (YGI_B_ECOLI); Fasta hit to ygiB_ECOLI (234 aa), 97% identity in 223 aa overlap. Contains a possible N-terminal signal sequence."
/codon_start=1
/transl_table=11
/product="possible lipoprotein"
/protein_id="CAD07713.1"
/db_xref="GI:16504265"
/db_xref="UniProt/TREMBL:Q8XGZ1"
/transl_table=11
/translation="MKRTKSIHHSFPRKMSABHLTPVALAVTAVFMLACGKSDTV SIYONADDCSAAHPGKSAECTTAYNNALKEAERAPKATAREDCVAERGGCCQAPR OAGAPENQAQAOQSSGSPMPLAGTMRGLMGKGGFAQDPLFSSKNPASPAYGKY TDAAGKRYGAAPGRTMTVPTKMAPAPATTVTYTRGFGESVAKOSTMORSAAGTST RSMGG"
1835. 1867
/gene="STY3365"
/note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
2434. 3597
/gene="STY3366"
2434. 3597
/gene="STY3366"
/note="Fasta hit to yrfC_ECOLI (387 aa), 50% identity in 392 aa overlap Orthologue of E. coli ygiC (YGI_C_ECOLI); Fasta hit to ygiC_ECOLI (386 aa), 94% identity in 386 aa overlap"
/codon_start=1
/transl_table=11

gene
1835. 1867
/note="conserved hypothetical protein"
/protein_id="CAD07714.1"
/db_xref="GI:16504266"
/db_xref="InterPro:IPR005494"
/db_xref="UniProt/TREMBL:Q8XRP6"
/transl_table=11
/translation="MERVSIITERPDRDRALEGFNFHTMYGEPVYCEDAYYKLTAAQ VKELSDVTAELHOMKLVKVRVYASDELMTKFLIPGHTGFWYQSQWQOOPSLRYLD LMDGIGEPKLENNADPTSLYEAFQPMIMEDQINMGPEGSQDONSLOEKLIB RPALEQYGFOLHLTCRDVYEDSGTQYODCAAEALNTEPYIDIDIGRGGO FTDQOVIANLFLYPMERMLKEMSTKEDAGVWLEPAKSTISNRALLPLNEM PPDHNNLPAYFAEDHPMDKVVKPIFSREAGANVSIENGKTIISVEGPEEEMT VQQPYLPKRGDSYTLIGSWLINDQAGIGIREDRALITQDSRFYPLFVEG"
complement(3659. 4489)
/gene="STY3367"
complement(3659. 4489)
/gene="STY3367"
/note="Orthologue of E. coli ygiD (YGI_D_ECOLI); Fasta hit to ygiD_ECOLI (271 aa), 87% identity in 266 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD07715.1"
/db_xref="GI:16504267"
/db_xref="GOA:Q823N7"
/db_xref="InterPro:IPR004183"
/db_xref="UniProt/TREMBL:Q823N7"
/transl_table=11
/translation="NMLTKLCSKDNHSLTGMPLFLGHGSPNVLDDNDYRAWR RIGBALRPOALVVSAMHYTRGTGTVAHERPDTLDFGSPALYDMHDPAGSPAL AGRVLVELLAPVVALDKBAMGPDHSGMGLIKKYPNADIPMQLSYDSTKPAWHPEM AGRLVALRDGCVLVAAGNVVNLRTVRMGDNI PYPMAASFNDPVKANTMGPEYEO HPLVNTLQHEGALSNPTEHFLPLLYVLGAMGKEPITIPVDIGEMGISMSLVQG "
4588. 5361
/gene="STY3368"
4588. 5361
/gene="STY3368"
/note="Orthologue of E. coli ygiE (YGI_E_ECOLI); Fasta hit to ygiE_ECOLI (257 aa), 93% identity in 257 aa overlap. Contains multiple possible membrane spanning hydrophobic domains"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD07716.1"
/db_xref="GI:16504268"
/db_xref="GOA:P67471"
/db_xref="InterPro:IPR003689"
/db_xref="UniProt/Swiss-Prot:P67471"
/transl_table=11
/translation="MSVPLILTLAAGATFGILGVLDQKPSNRVLAISLGAAGIM LLSLMEMLPALDTRGMSPVYGGMFTIGLGYRGLDRLLPRAHPQDLYOKROOPLP GSIKRTAILTLGTSLSLHNPBEGATPTVASSNLELGFALAVALLNINTEVGAAPV YAVYGSKRRTAIPWAGISMAELIGVLAALLIGSLVSPVMAIMAAVAGIVALSVD EMLPLAKRIDPNNNPISYGVLCGMSIMGLSVILQITIGIG"
6004. 7800
/gene="STY3370"
6004. 7800
/note="synonym: aasT"
/gene="STY3370"
/note="Similar to Klebsiella pneumoniae ariA sulfate sulfotransferase aasT TR:P97036 (EMBL:U32616) (598 aa) fasta scores: E(): 0, 86.84 id in 598 aa"
/codon_start=1
/transl_table=11
/product="probable ariA sulfate sulfotransferase"
/protein_id="CAD07717.1"
/db_xref="GI:16504269"
/db_xref="GOA:Q823N6"
/db_xref="InterPro:IPR010262"
/db_xref="UniProt/TREMBL:Q823N6"
/transl_table=11
/translation="MPDQYRKRTIAGVAVLTCGLTAASFPAGFOPAGKIGAVV DPFGNAPLVALVELDLSIDVAVYVAGGKGVPTVYVGSLESTYDGIPLFGYQ KFLANVTVEYKENGKMKDDYVQTSALVNHVMDNSISDLQOTKVIKVAPEGDELY LVVTHHTFQGAEFHNGEKDKNAGIILDAGPAGALPFDIAPTFVVDVQGEYRWILD

```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 20:04:03 ; Search time 550 Seconds  
(without alignments)  
6079.393 Million cell updates/sec

Title: US-10-520-820-28

Perfect score: 1434  
Sequence: 1 atgaagaatgaacgtgcgcaga.....aacagataaataaagctaa 1434

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing filter 45 summaries

Database :

Published Applications NA New:  
1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	1434	US-10-520-820-28	Sequence 28, App1
2	259.4	18.1	377	US-11-096-686-1544	Sequence 1812, Ap
3	253.4	17.7	960	US-10-467-657-1253	Sequence 3253, Ap
4	108.2	7.5	507	US-10-467-657-547	Sequence 547, App
5	64.6	4.5	510	US-11-098-686-8919	Sequence 8919, Ap
6	64.6	4.5	1457619	US-11-098-686-8919	Sequence 8919, Ap
7	62.4	4.4	1038	US-11-098-686-9716	Sequence 9716, Ap
8	62.4	4.4	1457619	US-11-098-686-8739	Sequence 8739, Ap
9	54.6	3.8	200	US-11-098-686-6509	Sequence 6509, Ap
10	43.2	3.0	3477	US-10-858-730-441	Sequence 141, App
11	37.2	2.6	1140	US-11-096-568A-3110	Sequence 3120, A
12	37.2	2.6	2253	US-11-096-568A-31087	Sequence 31087, A
13	37	2.6	88421	US-11-205-109-1	Sequence 1, App1
14	36.6	2.6	699	US-10-467-657-7053	Sequence 7053, Ap
15	36.6	2.6	3513	US-10-467-657-7841	Sequence 7841, Ap
16	36.4	2.5	1349	US-11-096-568A-26296	Sequence 26296, A
17	34.8	2.4	200	US-11-098-686-1545	Sequence 1545, Ap
18	34.6	2.4	1412	US-11-096-568A-10264	Sequence 10264, A
19	34	2.4	789	US-10-467-657-4617	Sequence 4617, Ap
20	33.8	2.4	1339	US-11-096-568A-6279	Sequence 6279, Ap

C	21	33.6	2.3	200	12	US-11-098-686-1544	Sequence 1544, Ap
C	22	33.4	2.3	200	12	US-11-098-686-6513	Sequence 6513, Ap
C	23	33.4	2.3	1632	9	US-11-096-568A-12879	Sequence 12879, A
C	24	33.4	2.3	1957	8	US-10-750-185-42527	Sequence 42527, A
C	25	33.4	2.3	1957	8	US-10-750-623-42527	Sequence 42527, A
C	26	33.2	2.3	1515	9	US-11-179-363-4	Sequence 4, App1
C	27	33.2	2.3	18394	8	US-10-995-561-13367	Sequence 13367, A
C	28	33.2	2.3	400	8	US-10-953-392-1	Sequence 17, App1
C	29	33	2.3	400	8	US-10-954-147-17	Sequence 17, App1
C	30	33	2.3	2196	12	US-11-052-554A-539	Sequence 539, App
C	31	32.8	2.3	813	8	US-10-467-657-2265	Sequence 2265, Ap
C	32	32.8	2.3	1551	9	US-11-096-568A-12361	Sequence 12361, A
C	33	32.8	2.3	2465	9	US-11-072-512-1278	Sequence 1278, Ap
C	34	32.6	2.3	678	8	US-11-187-622-14	Sequence 14, App1
C	35	32.6	2.3	879	8	US-10-965-972-5	Sequence 5, App1
C	36	32.6	2.3	1161	9	US-11-096-568A-23715	Sequence 23715, A
C	37	32.6	2.3	1165	9	US-11-096-568A-20738	Sequence 20738, A
C	38	32.6	2.3	5071	12	US-11-090-617-563	Sequence 563, App
C	39	32.6	2.3	76427	12	US-11-124-367A-5041	Sequence 5041, App
C	40	32.6	2.3	415117	8	US-10-995-561-13374	Sequence 13374, A
C	41	32.4	2.3	432	8	US-10-467-657-3275	Sequence 3275, Ap
C	42	32.4	2.3	438	8	US-10-953-392-14	Sequence 14, App1
C	43	32.4	2.3	438	8	US-10-954-147-14	Sequence 14, App1
C	44	32.4	2.3	1213	7	US-10-537-897-18	Sequence 18, App1
C	45	32.4	2.3	1213	7	US-10-537-897-36	Sequence 36, App1

#### ALIGNMENTS

RESULT 1	US-10-520-820-28	Application US/10520820
Sequence 28, Application US/10520820	US-10-520-820-28	US20060003393A1
GENERAL INFORMATION:		
APPLICANT: MUTABILIS S.A.		
TITLE OF INVENTION: Pathogenicity determinants which can be used as targets for develop		
TITLE OF INVENTION: means for preventing and controlling bacterial infections and/or		
TITLE OF INVENTION: disinfection		
FILE REFERENCE: 1621		
CURRENT APPLICATION NUMBER: US/10/520, 820		
CURRENT FILING DATE: 2005-01-07		
PRIOR APPLICATION NUMBER: PCT/EP2003/008209		
PRIOR FILING DATE: 2003-07-09		
PRIOR APPLICATION NUMBER: FR 0208636		
PRIOR FILING DATE: 2002-07-09		
NUMBER OF SEQ ID NOS: 32		
SOFTWARE: PatentIn version 3.1		
SEQ ID NO 28		
LENGTH: 1434		
TYPE: DNA		
ORGANISM: Escherichia coli		
US-10-520-820-28		
Query Match	100.0%	Score 1434; DB 8; Length 1434;
Best Local Similarity	100.0%	Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGAAAGTAAACGTCGACAGATTGAACGTGACAGAGTATGTTGGTGTGATGATG 60
DB	1	ATGAAAGTAAACGTCGACAGATTGAACGTGACAGAGTATGTTGGTGTGATGATG 60
QY	61	CTGAGATCTTACCTGTCGACGCGCCACCACTGATTCGCCGGAACGCCGCGCGTG 120
DB	61	CTGAGATCTTACCTGTCGACGCGCCACCACTGATTCGCCGGAACGCCGCGCGTG 120
QY	121	GTTAAAGTAAATCATGCAAGACGTCGCGGCGCGCGCTTAACTGCGATGATATC 180
DB	121	GTTAAAGTAAATCATGCAAGACGTCGCGGCGCGCGCTTAACTGCGATGATATC 180
QY	181	GCTTCTCTGCTGATCATGACGCGCTGCTGCGGTTGACGCGGATTCATGATGATC 240
DB	181	GCTTCTCTGCTGATCATGACGCGCTGCTGCGGTTGACGCGGATTCATGATGATC 240

QY 241 GCGGTGAGTAAATCTGCGCCGAGCTCAAGCTCAATGCGCACTTGTTCTGTACCGAG 300  
 DB 241 GCGGTGAGTAAATCTGCGCCGAGCTCAAGCTCAATGCGCACTTGTTCTGTACCGAG 300  
 QY 301 CATCCGACATTAACCAATTAACGGGTATCTTCCGCAACCAAGCTGATCCGTCTGAT 360  
 DB 301 CATCCGACATTAACCAATTAACGGGTATCTTCCGCAACCAAGCTGATCCGTCTGAT 360  
 QY 361 TTGGAAGAGTTCGAAAGGTGATTCGCGAGCGCTGCAAGCGGATTAACAGAGG 420  
 DB 361 TTGGAAGAGTTCGAAAGGTGATTCGCGAGCGCTGCAAGCGGATTAACAGAGG 420  
 QY 421 CTGAGTTGATTTGCGCGCTGCTTCTGTAACGCAAGGATGCGTGGCAAGGCTA 480  
 DB 421 CTGAGTTGATTTGCGCGCTGCTTCTGTAACGCAAGGATGCGTGGCAAGGCTA 480  
 QY 481 CAGCAGATGATCAACTGCGCGTAAAGCGGCTGCTGATGATCCAAAGGT 540  
 DB 481 CAGCAGATGATCAACTGCGCGTAAAGCGGCTGCTGATGATCCAAAGGT 540  
 QY 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAAGCGGATTCCTCGGAATTGGA 600  
 DB 541 ACCGATTTTGAAGCGCTACCGCGCGCTACGCTGTTAAAGCGGATTCCTCGGAATTGGA 600  
 QY 601 GCTGTGTCGCTAAATGTAAGACCGAAGAGATTTGTTAGCGCGCATGAATCTGAT 660  
 DB 601 GCTGTGTCGCTAAATGTAAGACCGAAGAGATTTGTTAGCGCGCATGAATCTGAT 660  
 QY 661 GCCGATTAACGAACTCTGCTGCTGTTAGTGAACCGGTTCCGAACGAGGTAAAGTCCCTGCTG 720  
 DB 661 GCCGATTAACGAACTCTGCTGCTGTTAGTGAACCGGTTCCGAACGAGGTAAAGTCCCTGCTG 720  
 QY 721 CAACCGGCTAAAGCGCGCTGCAATAGCAACCAAGCGCGAAGGTAAAGGTTAAC 780  
 DB 721 CAACCGGCTAAAGCGCGCTGCAATAGCAACCAAGCGCGAAGGTAAAGGTTAAC 780  
 QY 781 GGTGCGGCGCAGACGCTGATTTGCGCTGCTGCGCAACGCTGCGAGCGGATTAATTCGCTG 840  
 DB 781 GGTGCGGCGCAGACGCTGATTTGCGCTGCTGCGCAACGCTGCGAGCGGATTAATTCGCTG 840  
 QY 841 GAAGAAGCTGCTCTTTTCCAAATGCGCGCTGCTGCTGCTGCGCAACGCTGCGAG 900  
 DB 841 GAAGAAGCTGCTCTTTTCCAAATGCGCGCTGCTGCTGCTGCGCAACGCTGCGAG 900  
 QY 901 TTCCAGCTTTGCGCGATGCAAGCTGGAATGCTGATGCTGGAAGCTGGAATGCTGGA 960  
 DB 901 TTCCAGCTTTGCGCGATGCAAGCTGGAATGCTGATGCTGGAAGCTGGAATGCTGGA 960  
 QY 961 GCGGTGATGACCGAAGAGAACTGAAGCTGCGCTGAGCGCGAGCGCTAAACGCTGTGAA 1020  
 DB 961 GCGGTGATGACCGAAGAGAACTGAAGCTGCGCTGAGCGCGAGCGCTAAACGCTGTGAA 1020  
 QY 1021 AAAGTGTGATGACCAACGCTGCTTTGCAATCTGCAAGCGCGAGCAAGCTCTTTATCTG 1080  
 DB 1021 AAAGTGTGATGACCAACGCTGCTTTGCAATCTGCAAGCGCGAGCAAGCTCTTTATCTG 1080  
 QY 1081 GCAAAATGCGCGCAAGCTGCTGAGCGCTGATGCTGCTGCAACAGGATGCTCAACC 1140  
 DB 1081 GCAAAATGCGCGCAAGCTGCTGAGCGCTGATGCTGCTGCAACAGGATGCTCAACC 1140  
 QY 1141 AAACGCTGAAAGGAGATTTCCCGCCCGGTAAACCACTGCAACAGCTATGATTTGCTG 1200  
 DB 1141 AAACGCTGAAAGGAGATTTCCCGCCCGGTAAACCACTGCAACAGCTATGATTTGCTG 1200  
 QY 1201 GGGCGACTGGAAGCGCTGCACTGCTGCTGTTGAAGAGACAGCGCGCAGCGCTTG 1260  
 DB 1201 GGGCGACTGGAAGCGCTGCACTGCTGCTGTTGAAGAGACAGCGCGCAGCGCTTG 1260  
 QY 1261 ATGCGCGGATCTTGGCAGATCTGCTGCTGTAAGCGCGCGCTTAACCAAGAGATTT 1320  
 DB 1261 ATGCGCGGATCTTGGCAGATCTGCTGCTGTAAGCGCGCGCTTAACCAAGAGATTT 1320

QY 1321 GCGGAGTAAAGAGTCTGCGCCAGCGTGGCGAAGTGTGCTCAACTTTGAAGAC 1380  
 DB 1321 GCGGAGTAAAGAGTCTGCGCCAGCGTGGCGAAGTGTGCTCAACTTTGAAGAC 1380  
 QY 1381 GGTGCTCGACGACCAATCATCAAGAGATTCACAGATTAAGGCTTA 1434  
 DB 1381 GGTGCTCGACGACCAATCATCAAGAGATTCACAGATTAAGGCTTA 1434

RESULT 2  
 US-11-116-881A-1812  
 ; Sequence 1812, Application US/11116881A  
 ; Publication No. US20060041949A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Dongmei  
 ; APPLICANT: Nielsen, Mark T.  
 ; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof  
 ; FILE REFERENCE: 07678/141014  
 ; CURRENT APPLICATION NUMBER: US/11/116,881A  
 ; CURRENT FILING DATE: 2005-04-27  
 ; PRIOR APPLICATION NUMBER: 60/665,451  
 ; PRIOR FILING DATE: 2005-03-24  
 ; PRIOR APPLICATION NUMBER: 60/665,097  
 ; PRIOR FILING DATE: 2005-03-24  
 ; PRIOR APPLICATION NUMBER: 60/646,764  
 ; PRIOR FILING DATE: 2005-01-25  
 ; PRIOR APPLICATION NUMBER: 60/607,357  
 ; PRIOR FILING DATE: 2004-09-03  
 ; PRIOR APPLICATION NUMBER: 60/566,235  
 ; PRIOR FILING DATE: 2004-04-29  
 ; PRIOR APPLICATION NUMBER: 10/934,944  
 ; PRIOR FILING DATE: 2004-09-03  
 ; PRIOR APPLICATION NUMBER: 10/943,507  
 ; PRIOR FILING DATE: 2004-09-17  
 ; PRIOR APPLICATION NUMBER: 60/503,989  
 ; PRIOR FILING DATE: 2003-09-18  
 ; PRIOR APPLICATION NUMBER: 60/485,368  
 ; PRIOR FILING DATE: 2003-07-08  
 ; PRIOR APPLICATION NUMBER: 60/418,933  
 ; PRIOR FILING DATE: 2002-10-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 2300  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 1812  
 ; LENGTH: 377  
 ; TYPE: DNA  
 ; ORGANISM: Nicotiana tabacum  
 ; US-11-116-881A-1812

Query Match 18.1%; Score 259.4; DB 9; Length 377;  
 Best Local Similarity 99.6%; Pred. No. 3.6e-70;  
 Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1174 CCACCTGAACAGGATGATTTGCTGCTGCGCACCTGGAAGCGGTGCACTGCGGATGCTG 1233  
 DB 1 1 CCACTGGAACAGGATGATTTGCTGCTGCGCACCTGGAAGCGGTGCACTGCGGATGCTG 60  
 QY 1234 TTGGAAGAGACAGCGCGAGCGCTGATGCGCGGATCTTGGCAGATCTGCTGTGAAA 1293  
 DB 61 TTGGAAGAGACAGCGCGAGCGCTGATGCGCGGATCTTGGCAGATCTGCTGTGAAA 120  
 QY 1294 GCGGCGACTATTAACCAAGAGATTTGCGGAGTAAAGAGCTTGGGCAACGCTGCG 1353  
 DB 121 GCGGCGACTATTAACCAAGAGATTTGCGGAGTAAAGAGCTTGGGCAACGCTGCG 180  
 QY 1354 GAAGTGTGCTCAACTTTGAAGCGGTTGCTGCAACCAATCATCAAGAGATC 1413  
 DB 181 GAAGTGTGCTCAACTTTGAAGCGGTTGCTGCAACCAATCATCAAGAGATC 240  
 QY 1414 CAACAGATTAAGGCTTA 1434  
 DB 241 CAACAGATTAAGGCTTA 261

Oy	682	CTGTTAGAGACCCGTTCCGAACAGGAGTATGTGCTGCTGCACACCGGGTAAACGCGCGCTG	741
Db	700	GTTTTCCTGACCCCGAGCGAAGAGGGCATGACCTCTGTTCAAGGAAGCGCAA---CCCAT	756
Oy	742	CATATGCCCAACCCAAAGCGCAGAGAACTATATAGCTTACCGGTGCGGCGACACGGTGATT	801
Db	757	TACACAGCCCAACCCGCGCCCAAGAAATTATACGACGTGTCGCGCGCAGCGACACCGTCATT	816
Oy	802	GGCGTCCCTGCGCGGCAACGCTGCGACGCGGTAAATTGCTGCGAAGAGCCCTCTTCTTGCGC	861
Db	817	GCCGGAATGGGCTTTGGGGCTGCGCGCAAGCTGCGACCATGCCCGAAGCCCATGATCCTTGCGC	876
Oy	862	AATCGCGCGCGCTGGCGTGTGATGATCGGCAAACTGAGAACCTCCACCGTGTTCGCCGATCGAG	921
Db	877	AATATCGCGGCGCGGGGTGTGCTGCGCAAACTCGGTAGCGCGCTTGTCTCGTTGTGAGAG	936
Oy	922	CTGGAATAATGC	932
Db	937	TTGCGTTGAAAGC	947

```

RESULT 4
US-10-467-657-547
Sequence 547: Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: Seqwin99, version 1.04
SEQ ID NO 547
LENGTH: 507
TYPE: DNA
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-547

Query Match          7.5%: Score 108.2; DB 8; Length 507;
Query Local Similarity 56.5%: Pred. No. 3e-23;
Matches 222; Conservative 0; Mismatches 168; Indels 3; Gaps 1

QY      1025 TGTGATGACCAACGCTGTCTTTGACATCTGACGCGCGGACGCTCTTATCTGGCA 108
DB      95  TGTGTTTACCAACGGCTGTTTGACATCTCCACAGGGGCGACGTTATCTGCGGC 154

QY      1085 ATGCCCCGAACTGGGTGACCGCTTGATTTGTGCCGTCAACGCGATCCTCCACCAAC 1144
DB      155 AGGCGCGCTTCGGCAGGGGCGCGCGTTGGTGCTGGCGTTGAATACGATCTTCGTGCGGC 214

QY      1145 GGCCTG---AAAGGGGATTCCTCCGCCCGGTAACCCACTCGAACAGGTATGATTTGCTGCG 1207
DB      215 GTTTGGGCAAGGGGGGGGCGACCGCCCGGTAAATCTTTGGAGAACCGTCCCGCGCTTGGCG 274

QY      1202 GCGCAGCTGGAAGCGGTGCGACTGGGTAGTGTGTTTGAAGAGGACACGCGCGAGCGCTTGA 1267
DB      275 CCGCGTTGGAAAGCTGTGATTTGTTGTTACTTGATTTGACAGAGATACGCGCGGGGGCTTGA 334

QY      1262 TCGCGCGGATCTTGCCAGATCTGCTGTGTGAAGGCGGCGACATATMAACAGAAAGATTG 1322
DB      335 TTGAGCGCGTCAACCGCGGAGGTTGGTCAAGGCGGCGGATTGGGTCGTGATTAAGATTG 394

QY      1322 CCGGAGATTAAGAGTCTGCGGCCCAACGCTGCGGAAGTTTGTGTCTCACTTTGAAGACG 1388
DB      395 TCGGTGCGCGGAAACCGCTGCGACGCGCGGCTGAGGTGTTTCAATCCGTTTCTGCAATC 454

```

Qy 1382 GTTCTGACGACCAACATCATCAAGATCC 1414  
 Db 455 AGACTTGACACCAAGACTTGGCAAAATCC 487

## RESULT 5

US-11-098-686-8919  
 ; Sequence 8919, Application US/11098686  
 ; Publication No. US20060024696A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
 ; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING  
 ; FILE REFERENCE: 09531-128001  
 ; CURRENT APPLICATION NUMBER: US/11/098,686  
 ; CURRENT FILING DATE: 2005-04-04  
 ; PRIOR APPLICATION NUMBER: PCT/US03/31318  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: US 60/416,395  
 ; PRIOR FILING DATE: 2002-10-04  
 ; NUMBER OF SEQ ID NOS: 11433  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 8919  
 ; LENGTH: 510  
 ; TYPE: DNA  
 ; ORGANISM: Lawsonia intracellularis  
 US-11-098-686-8919

Query Match 4.5%; Score 64.6; DB 12; Length 510;  
 Best Local Similarity 48.9%; Pred. No. 1e-09;  
 Matches 203; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

Qy 1014 TGGTGAAGAAAGTGTGATGACCAACGCTCTTGTGACATCTCGACGCGGCGACGCTC 1073  
 Db 81 TGGGAAAAAATTAATCTTACAAACGGTTGTATGATATCTTCACTCGACATCTTGA 140  
 Qy 1074 TTATCTGGCAATGCGCCGCAAGCTGGGTGACCGCTTGATGTCCTGCAACAGATGC 1133  
 Db 141 TCTTCTGACTAGAGAAAAACATATGAGATATCTGCTTGGATTAATACTGATTA 200  
 Qy 1134 CTCACCAACGCTGGAAGAGGATCCC--GCCCGTAACCCACTCGAACGCTAT 1190  
 Db 201 CTCGTAAACGGTTAGAGAAAAAGATCCAGATGACATTTATCTTCCATATTCGTGC 260  
 Qy 1191 GATGTGCTGGCGCACTGGAAGCGGTGACCTGGGTAGTGTGTTGAAGAGACGCGC 1250  
 Db 261 TTTTGTATTACCTCATCTTGAACCTTATGATTTCAATTAATGTTTGAAGAGATACCC 320  
 Qy 1251 GCAGCGCTTGATCGCCGGATCTTGCCAGATCTGCTGTGAAGCGCGCACTATAACC 1310  
 Db 321 ATTACAATTATTAAGCTATTTCAACCTAATGTACTGTTAAAGGTGTGATGTGTGT 380  
 Qy 1311 AGAAGAGTTGCCGGAGTAAAGAGTGTGGCCCAAGCTGGCGAAGTGTGCTCA 1370  
 Db 381 AGAACAAATTAATAGTAAAGGTTGTAGACGCGCATATGCTTGTATTAAGTCTTCC 440  
 Qy 1371 CTTTGAAGACGCTGCTGACGACCAACATCATCAAGAAATCCAAAGATTA 1425  
 Db 441 TTTTCTTGAAGACATTAACCAAGTGGTGTAGTACAAAATAAGTATTAATAA 495

## RESULT 6

US-11-098-686-8739/C  
 ; Sequence 8739, Application US/11098686  
 ; Publication No. US20060024696A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
 ; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING  
 ; FILE REFERENCE: 09531-128001  
 ; CURRENT APPLICATION NUMBER: US/11/098,686  
 ; CURRENT FILING DATE: 2005-04-04  
 ; PRIOR APPLICATION NUMBER: PCT/US03/31318

; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: US 60/416,395  
 ; PRIOR FILING DATE: 2002-10-04  
 ; NUMBER OF SEQ ID NOS: 11433  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 8739  
 ; LENGTH: 1457619  
 ; TYPE: DNA  
 ; ORGANISM: Lawsonia intracellularis  
 US-11-098-686-8739

Query Match 4.5%; Score 64.6; DB 12; Length 1457619;  
 Best Local Similarity 48.9%; Pred. No. 1.4e-08;  
 Matches 203; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

Qy 1014 TGGTGAAGAAAGTGTGATGACCAACGCTCTTGTGACATCTCGACGCGGCGACGCTC 1073  
 Db 21403 TGGGAAAAAATTAATCTTACAAACGGTTGTATGATATCTTCACTCGACATCTTGA 21344  
 Qy 1074 TTATCTGGCAATGCGCCGCAAGCTGGGTGACCGCTTGATGTCCTGCAACAGATGC 1133  
 Db 21343 TCTTCTGACTAGAGCAAAAACATATGAGATATCTGCTTGGATTAATACTGATTA 21284  
 Qy 1134 CTCACCAACGCTGGAAGAGGATCCC--GCCCGTAACCCACTCGAACGCTAT 1190  
 Db 21283 CTCGTAAACGGTTAGAAAAAGATCCAGATGACATTTATCTTCCATATTCGTGC 21224  
 Qy 1191 GATGTGCTGGCGCACTGGAAGCGGTGACCTGGGTAGTGTGTTGAAGAGACGCGC 1250  
 Db 21223 TTTTGTATTACCTCATCTTGAACCTTATGATTTCAATTAATGTTTGAAGAGATACCC 21164  
 Qy 1251 GCAGCGCTTGATCGCCGGATCTTGCCAGATCTGCTGTGAAGCGCGCACTATAACC 1310  
 Db 21163 ATTACAATTATTAAGCTATTTCAACCTAATGTACTGTTAAAGGTGTGATGTGTGT 21104  
 Qy 1311 AGAAGAGTTGCCGGAGTAAAGAGTGTGGCCCAAGCTGGCGAAGTGTGCTCA 1370  
 Db 21103 AGAACAAATTAATAGTAAAGGTTGTAGACGCGCAATTAATGCTTGTATTAAGTCTTCC 21044  
 Qy 1371 CTTTGAAGACGCTGCTGACGACCAACATCATCAAGAAAGATCCAAAGATTA 1425  
 Db 21043 TTTTCTTGAAGACATTAACCAAGTGGTGTAGTACAAAATAAGTATTAATAA 20989

## RESULT 7

US-11-098-686-9716  
 ; Sequence 9716, Application US/11098686  
 ; Publication No. US20060024696A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
 ; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING  
 ; FILE REFERENCE: 09531-128001  
 ; CURRENT APPLICATION NUMBER: US/11/098,686  
 ; CURRENT FILING DATE: 2005-04-04  
 ; PRIOR APPLICATION NUMBER: PCT/US03/31318  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: US 60/416,395  
 ; PRIOR FILING DATE: 2002-10-04  
 ; NUMBER OF SEQ ID NOS: 11433  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 9716  
 ; LENGTH: 1038  
 ; TYPE: DNA  
 ; ORGANISM: Lawsonia intracellularis  
 US-11-098-686-9716

Query Match 4.4%; Score 62.4; DB 12; Length 1038;  
 Best Local Similarity 49.1%; Pred. No. 6.1e-09;  
 Matches 165; Conservative 0; Mismatches 171; Indels 0; Gaps 0;  
 Qy 37 GTGATGTGTTGGTGTGATGATGCTGATCGTTACTGTTACGCCCCACCAAGCTGATC 96



Db	67	GTCTTGTGCTTGGTGATATTAATGTAATTAATTAATTTATGAGTTTCAGATCGATT	126
Qy	97	TTCCGCGGAAGCCGCGGTGCGCGGTAAAGTAATACCATCGAAGAACGTCGCGGCGC	156
Db	127	TCTCTGAGGACCTCTCCCTATTGTAATAATTAAGAAATGAAGAACATCTCTTGCGGGGT	186
Qy	157	GCGGCTPAACGTGCGATGAAATATCGCTTCTCTCGGTGTAAATGCACGCTGSGGTG	216
Db	187	GCAGGAAATGTAAGCAAGAGTATTCGCTCTCGGCTGAAAGGTAAACGATTAATGGGGCT	246
Qy	217	AAGGGCATTAACGATACGACGCGCGCGCGCGAGTAATCTGCGCGCAAGCTCAAGTCAA	276
Db	247	GTAAGCAACACCAAGATGAGAAAAATTCAAGATCTTTAATGTAACAAGGGAATTTTA	306
Qy	277	TGCGATTCGTTTCTGTACGACGCAATCCGACATTACAATTAACGGTACTTTCGCG	336
Db	307	TCTTCATTTAATCATTTGCAATCGTCAACACAGTAATAAACACGATGATGGCTCAC	366
Qy	337	AACCAACAGTATCGCTGTGATTTTGAAGAGGT	372
Db	367	AGACACACAGTATACGACTGTATCAATGAGAGGT	402

```

RESULT 8
US-11-098-686-8739
Sequence 8739, Application US/11098686
Publication No. US20060024696A1
GENERAL INFORMATION:
APPLICANT: Kapur, Vivek and Gebhart, Connie J.
TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
FILE REFERENCE: 09531-128001
CURRENT APPLICATION NUMBER: US/11/098,686
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: PCT/US00/31318
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: US 60/416,395
PRIOR FILING DATE: 2002-10-04
NUMBER OF SEQ ID NOS: 11433
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 8739
LENGTH: 1457619
TYPE: DNA
ORGANISM: Lawsonia intracellularis
US-11-098-686-8739

```

	Query Match	4.4%;	Score 62.4;	DB 12;	Length 1457619;
	Best Local Similarity	49.1%;	Pred. No. 6,6e-08;		
	Matches 165;	Conservative	0;	Mismatches 171;	Indels 0;
				Gaps	0;
Qy	37	GTGATGGTGGTGGTGGATGATGATCGTATCGTTACTGTGACGGCCCCACGATCGTATC	96		
Db	1014025	GTTCTTTCGTTGCGTTGGATGATATATGTATGATATATATCTTATTTGGTGTTTCAGATCGTATTT	1014084		
Qy	97	TGCGCCGGAAGGCGCGGTGCGCGGTGTTAAAGTGAATACATCGAAGAACGTCCGGCGGCG	156		
Db	1014085	TCTCTTGAGGACCTGTCTCCATTGTAAAAATAGAGATGAAAAAGCAATCTCTTGGGGGCT	1014144		
Qy	157	GGCGCTAACGTCGGGAGTGAATATGTCCTCTCCGGTGTAAATGCAAGCGCTGTCGGGGTTG	216		
Db	1014145	GCAGGAGATGTAGCAAGGATATGTCTCTCTCGGTGAAAGGTBAACATTAATTTGGGCT	1014204		
Qy	217	ACGGGCAATTCAGATGCAAGCGCGCGCTGAGTAAATCTCTGCGCGAAGCTCAAGCTCAAA	276		
Db	1014205	GTAAGAACAAAGCAAAAGTGAGAAAATAATTCAAAGTCTTTTAATGACAAAGGGAATTTTA	1014264		
Qy	277	TGCGACTTCGTTTCTGTACCGAAGCATTCGACCATTAACCAAAATTAAGGGTACTTTCCCGC	336		
Db	1014265	TCTTCAATTAATTAACATTTCGCAATCGTCAAAACAACGTAAAAAACAAGTGTATGGCTCAC	1014324		
Qy	337	AACCAACAGCTGATCCGCTGATTTTGAAGAAAGT	372		
Db	1014325	AGACAAACGATGATACGACTTGATTCATGAGAAAGT	1014360		

RESULT 9  
 US-11-098-686-6509  
 Sequence 6509, Application US/11098686  
 Publication No. US20060024656A1  
 GENERAL INFORMATION:  
 APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
 TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
 TITLE OF INVENTION: FROM LAMSONIA INTRACELLULARIS AND METHODS OF USING  
 FILE REFERENCE: 09531-128001  
 CURRENT APPLICATION NUMBER: US/11/098,686  
 CURRENT FILING DATE: 2005-04-04  
 PRIOR APPLICATION NUMBER: PCT/US03/31318  
 PRIOR FILING DATE: 2003-10-01  
 PRIOR APPLICATION NUMBER: US 60/416,395  
 PRIOR FILING DATE: 2002-10-04  
 NUMBER OF SEQ ID NOS: 11433  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 6509  
 LENGTH: 200  
 TYPE: DNA  
 ORGANISM: Lamsonia intracellularis  
 US-11-098-686-6509

	Query Match	3.8%	Score 54.6	DB 12	Length 200
	Best Local Similarity	58.2%	Pred. No. 9.3e-07		
	Matches	96	Conservative	0	Mismatches 65
					Indels 0
					Gaps 0
QY	37	GTGATGTGTGTGGTGTGATGTGATCGTACTGTGTAACGGCCACCAAGTGTATC	96		
DB	25	GTTCTGTTCGTGTGTGATATATGTATGATATATCTATTTGGTGTTCAGATGTAT	84		
QY	97	TGCGCCGAAGCGCCGGTCCCGGTGTTAAAGTAATACATCGAAGAACGTCCGGCGCC	156		
DB	85	TCTCTTGAGGCACTGTCTCCATTGTAAAAATAGAGATGAAAGCAATCTCTTGGGGGT	144		
QY	157	GCGGCTAACGTGGGAGTAATATGGCTCTCTCGGTGTATATGCA	201		
DB	145	GCAAGGAATGTAGCAAGAGTATGTCTCTCTGGGTGAAAGCTA	189		

```

RESULT 10 730-141
US-10-858-730-141
Sequence 141, Application US/10858730
Publication No. US2005025556A1
GENERAL INFORMATION:
APPLICANT: Bailey, Richard B.
APPLICANT: Blomquist, Paul
APPLICANT: Doten, Reed
APPLICANT: Driggers, Edward M.
APPLICANT: Madden, Kevin T.
APPLICANT: O'Leary, Jessica
APPLICANT: O'Toole, George
APPLICANT: Trueheart, Joshua
APPLICANT: Walbridge, Michael J.
APPLICANT: Yorgey, Peter S.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
TITLE OF INVENTION: PRODUCTION
FILE REFERENCE: 14184-030001
CURRENT APPLICATION NUMBER: US/10/858, 730
CURRENT FILING DATE: 2004-06-01
PRIOR APPLICATION NUMBER: US 60/475, 000
PRIOR FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: US 60/551, 860
PRIOR FILING DATE: 2004-03-10
NUMBER OF SEQ ID NOS: 364
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 141
LENGTH: 3477
TYPE: DNA
ORGANISM: Thermobifida fusca
US-10-858-730-141

```

Query Match 3.0%; Score 43.2; DB 8; Length 3477;  
Best Local Similarity 46.2%; Pred. No. 0.0082;  
Matches 144; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 1053 CTTGACGCGCGGACGCTCTTATCTGCGAATGCCCGACAGTGGGTGACCGCTTGAT 1112  
DB 1131 CATGCTGACCTGTGCGCTGCACTATGTGGACGCGAGCGGGGTGCGCATATGCGGAGCT 1190  
QY 1113 TGTGCGGTCAACAGCGATGCTCCACCAACCGCTGAAGGGATTCGCCGCCGTTAA 1172  
DB 1191 GGCCTCCCGGTGCGACCGCTCCACGCTGCGTGTGACTGCTCCAGAGTAGC 1250  
QY 1173 CCCACTCGAAGAGGTATGATTTGCTGGGCGCACTGGAACCGGTGCACTGGATGTC 1232  
DB 1251 GGTACTGGAAGCTGAGACTGAGATGCTGGGCGCGCGCTGCTCAACTCGTCACTA 1310  
QY 1233 GTTTGAAGAGCAACGCGCGACGCTTGATGCGCGGATCTTGGCAGATCTGTGTAA 1292  
DB 1311 CGAGGACGCGCGACCGCTGACTCCCGTTGCGCAAGGTGCGCGCTGGGCGTGAACA 1370  
QY 1293 AGGCGGCGACTATTAACGAGAGATTTCCGGGAGTAAAGATCTGGGCGAACGGTGG 1352  
DB 1371 CGGGCGCGCCTCTCATGTGGCGTGAACATGACAGACAGGCGCGGACCGCGAAGC 1430  
QY 1353 CGAAGTGTGGT 1364  
DB 1431 GAAAGTGAAGT 1442

## RESULT 11

US-11-096-568A-31210  
; Sequence 31210, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 31210  
; LENGTH: 1140  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1140)  
; OTHER INFORMATION: Ceres Seq. ID no. 13579027  
US-11-096-568A-31210

Query Match 2.6%; Score 37.2; DB 9; Length 1140;  
Best Local Similarity 49.5%; Pred. No. 0.41;  
Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 679 GCTCTGTATGACCCGTTCCGAACAGGGTATGCTGCTGCAACCGGGTAAAGCCCG 738  
DB 862 GTTCTAGTAAACTGGGTCCAAAGATCTGACATTTATACAAAGGAAAAACAATC 921  
QY 739 CTGCATATGCCAACCGACGCGAGAAAGTATGACGTTACCGGTGGCGGACACGGTG 798  
DB 922 CAGCAGCTATCATATCCGGCTGCAACAAGTGTGTACTACAGAGGCTGGGATATCTTC 981  
QY 799 ATTGGCGTCTGGGCGCAACGCTGCGACGGGTAAATTGCTGGAAGAAGCTGCTTTT 858  
DB 982 ACTCAGCATTTGCAAGTGTGATGTAGAGGCAAGTCCCAATGAGAAATGCTGAATTT 1041  
QY 859 GCCAATGCGCGCG 872  
DB 1042 GCTGCTGCACTGC 1055

## RESULT 12

US-11-096-568A-31087  
; Sequence 31087, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 31087  
; LENGTH: 2253  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(2253)  
; OTHER INFORMATION: Ceres Seq. ID no. 4987027  
US-11-096-568A-31087

Query Match 2.6%; Score 37.2; DB 9; Length 2253;  
Best Local Similarity 49.5%; Pred. No. 0.52;  
Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 679 GCTCTGTATGACCCGTTCCGAACAGGGTATGCTGCTGCAACCGGGTAAAGCCCG 738  
DB 1975 GTTCTAGTAAACTGGGTCCAAAGATCTGACATTTATACAAAGGAAAAACAATC 2034  
QY 739 CTGCATATGCCAACCGACGCGAGAAAGTATGACGTTACCGGTGGCGGACACGGTG 798  
DB 2035 CACGAGTATCATATCCGGCTGCAACAAGTGTGATATACAGGAGCTGGGATATCTTG 2094  
QY 799 ATTGGCGTCTGGGCGCAACGCTGCGACGGGTAAATTGCTGGAAGAAGCTGCTTTT 858  
DB 2095 ACTCAGCATTTGCAAGTGTGATGTAGAGGCAAGTCCCAATGAGAAATGCTGAATTT 2154  
QY 859 GCCAATGCGCGCG 872  
DB 2155 GCTGCTGCACTGC 2168

## RESULT 13

US-11-205-109-1  
; Sequence 1, Application US/11205109  
; Publication No. US20050287641A1  
; GENERAL INFORMATION:  
; APPLICANT: Farnet, Chris  
; APPLICANT: Zazopoulos, Emmanuel  
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS  
; FILE REFERENCE: 3002-2US  
; CURRENT APPLICATION NUMBER: US/11/205,109  
; CURRENT FILING DATE: 2005-08-17  
; PRIOR APPLICATION NUMBER: US/09/976,059  
; PRIOR FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/239,924  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 88421  
; TYPE: DNA  
; ORGANISM: Actinoplanes sp.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (2077)..(3078)  
; OTHER INFORMATION: ORF 1; positive strandedness  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (3118)..(4032)  
; OTHER INFORMATION: ORF 2; positive strandedness



QY 341 AACACTGATCCGCTGATTTTGAAGAAGTTTGAAGGTGTTGATCCGACCGCTGC 400  
DB 63732 GCCGGCCGGTTCTGTGTGTGACAGCGTGATCTGCCGCGCGGACCGAAGCTGCGC 63791  
QY 401 ACGAGCGGATTTATCAGCGCTGAGTTGATGGCGGCTGGTCTTTCTGACTACCCA 460  
DB 63792 GGGTGGCGGCGGCTGCTGTATCGATGTACACTCGGGTTCACCGGCGCGGCCA 63851  
QY 461 AAGGTGGCTGGCAAGGCTACAGAGATTCGAAGTCCGCGCGTAAAGCGGGT 513  
DB 63852 AGGGTGTGTGACCAACCAACCAATCTGTGTGATCTTGGCAACGACGTTGT 63904

RESULT 14  
US-10-467-657-7053

; Sequence 7053, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWIn99, version 1.04  
; SEQ ID NO 7053  
; LENGTH: 699  
; TYPE: DNA  
; ORGANISM: Neisseria gonorrhoeae  
; US-10-467-657-7053

Query Match 2.6%; Score 36.6; DB 8; Length 699;  
Best Local Similarity 47.6%; Pred. No. 0.54;  
Matches 108; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 1061 CCGGACAGTCTTATCTGGAATGCGCGCAAGCTGGTGACCGCTTATGTTGCCG 1120  
DB 440 CCGGTCAAGTGTGCAATGCGGTAAAGCCGCAATGCGCGAGATGCGATTAATCTTCCG 439  
QY 1121 TCACACAGGATGCTTCCACCAACGCGTGAAGGGGATTCGCCCGGTAAACCACTCG 1180  
DB 500 TGTATGTCGACCAACCGCCCAACCGCGGCTTACTATATTATGATAAGAAAGCG 559  
QY 1181 AACAGCATATGATGCTGGGCGCACTGGAAGCGTCACTGGGTAGTGTCTTTGAAG 1240  
DB 560 ATGTGCGGAAGTCAATATGAGCGTGAAGCAAGCGTTGAATATGATTTGCTGGGTA 619  
QY 1241 AGAACAAGCGCGAGCGCTGATGCGCGGATCTTGCCAGATCTGCTG 1287  
DB 620 TGTATATCTTCAAGCACTGCGCTCAAAACATGTGCAAGCACTATG 666

RESULT 15  
US-10-467-657-7841

; Sequence 7841, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8

; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWIn99, version 1.04  
; SEQ ID NO 7841  
; LENGTH: 3513  
; TYPE: DNA  
; ORGANISM: Neisseria gonorrhoeae  
; US-10-467-657-7841

Query Match 2.6%; Score 36.6; DB 8; Length 3513;  
Best Local Similarity 45.6%; Pred. No. 0.91;  
Matches 129; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 553 CGCTACCGGCGCGCTAGCTGTAAAGCGCAATCTCGGAATTTGAAGCTGTGTCGGT 612  
DB 1588 CGCGATGTGGCGCAAGCGCAACCGATTAAACCTGCGGACCAACGCGGATGTGTGCGT 1647  
QY 613 AAATGTAAAGCGAAGAGATTTTGAAGCGCGGCAATGAATGCGGATTAAGAA 672  
DB 1648 ACGGTACGTTTCCGATGCGCGCTTGCCCAAGAACGGTGTGTGACAGCGCTTGGCGCG 1707  
QY 673 CTCTCGCTCTGTATGACCGGTTCCGAACAGGGTATGTGCTGTGCAACCGGGTAA 732  
DB 1708 TTCCCGAATGAGTGCACACCTGCGCGCAACGCGCGCTGCTGCGCGCTTTCGCC 1767  
QY 733 GCGCGCTGATATGCAACCAAGCGAGAAAGTATGACGTTACCGGTGCGGCGAC 792  
DB 1768 GACTGTGCAAGACGACACCGCGCGCTGATGATCTTGGCGGCAAGCGGCAAA 1827  
QY 793 ACGGTATGCGCTCTGCGGCGCAACGCTGGCAGCGGTAATT 835  
DB 1828 ACGGTAAACACCGCATTTGCCGAAGTGCAGAACCGTGCATT 1870

Search completed: March 18, 2006, 20:19:09  
Job time : 554 secs